SIAM J. MATRIX ANAL. APPL.
Vol. 43, No. 1, pp. 68–93

TWICE IS ENOUGH FOR DANGEROUS EIGENVALUES*

ANDREW HORNING† AND YUJI NAKATSUKASA‡

Abstract. We analyze the stability of a class of eigensolvers that target interior eigenvalues with rational filters. We show that subspace iteration with a rational filter is robust even when an eigenvalue is near a filter's pole. These dangerous eigenvalues contribute to large round-off errors in the first iteration but are self-correcting in later iterations. For matrices with orthogonal eigenvectors (e.g., real-symmetric or complex Hermitian), two iterations are enough to reduce round-off errors to the order of the unit round-off. In contrast, Krylov methods accelerated by rational filters with fixed poles typically fail to converge to unit round-off accuracy when an eigenvalue is close to a pole. In the context of Arnoldi with shift-and-invert enhancement, we demonstrate a simple restart strategy that recovers full precision in the target eigenpairs.

Key words. subspace iteration, Arnoldi, shift-and-invert, rational filters, FEAST, CIRR

AMS subject classifications. 65F15, 65G50, 15A18

DOI. 10.1137/20M1385330

1. Introduction.

When combined with shift-and-invert enhancement, subspace iteration and Arnoldi are two classic iterative schemes for computing a few interior eigenvalues of an $n \times n$ matrix A. Each method constructs an orthonormal basis for a search subspace by iteratively applying the spectral filter

$$s(A) = (zI - A)^{-1} \tag{1.1}$$

to a set of vectors. Approximate eigenpairs can then be extracted from the search subspace with a projection step, e.g., Rayleigh–Ritz. The shift z is selected to target a region of interest, and both methods typically approximate eigenvalues of A closest to z.

Recently, general rational filters of the form

$$r(A) = \sum_{j=1}^{\ell} \omega_j (z_j I - A)^{-1} \tag{1.2}$$

have attracted a great deal of attention in the context of large, data-sparse eigenvalue problems [1, 4, 7, 8, 11, 14, 17]. When the weights $\omega_1, \ldots, \omega_\ell$ and nodes $z_1, \ldots, z_\ell$ are chosen appropriately, these rational filters can robustly target eigenvalues in a region of interest and significantly accelerate the convergence of the subspaces constructed by subspace iteration, Arnoldi, or variants thereof [1, 17]. They also tend to be highly parallelizable because each shift-and-invert transformation may be applied independently [8].

In his 2001 volume on matrix algorithms for eigenvalue problems, Stewart noted that shift-and-invert Arnoldi encounters difficulties in floating-point arithmetic when

*Received by the editors December 11, 2020; accepted for publication (in revised form) September 1, 2021; published electronically January 11, 2022.
https://doi.org/10.1137/20M1385330
Funding: The work of the first author was partially supported by NSF DMS-1818757.
†Center for Applied Mathematics, Cornell University, Ithaca, NY 14853 USA (ajh326@cornell.edu).
‡Mathematical Institute, University of Oxford, Oxford, OX2 6GG UK (nakatsukasa@maths.ox.ac.uk).

the shift lies too close to an eigenvalue of A [15, p. 309]. Although the eigenvalue adjacent to the shift is rapidly approximated to the order of the unit round-off u, the residuals of other computed eigenpairs stagnate near the order of u/d, where d is the distance between the "dangerous" eigenvalue and the shift. This phenomenon has also recently been observed in the context of Krylov methods, where the subspace is constructed with contour integrals and rational approximation [1].

Curiously, dangerous eigenvalues do not inflict the same stagnation in the residuals of the other target eigenpairs during subspace iteration. Figure 1.1 compares the residuals of two target eigenpairs computed with Arnoldi (left) and subspace iteration (right) using the shift-and-invert filter in (1.1) with $z = 10$. The approximation to the dangerous eigenvalue $\lambda_1 = 10 + 10^{-12}$ converges rapidly to unit round-off accuracy in both cases. However, only subspace iteration computes an approximation to the second target eigenvalue $\lambda_2 \approx 10.1$ to unit round-off accuracy.

A similar story unfolds in Figure 1.2, where we compute two target eigenpairs with the contour integral eigensolver described in [17]; one eigenvalue is located at a distance of 10^{-10} from the contour. As we refine the quadrature along the contour, the poles of a rational filter with form (1.2) cluster near the dangerous eigenvalue, and we observe that the residual of the dangerous eigenpair converges rapidly to unit round-off, while the residuals of the remaining target pairs stagnate near 10^{-5}. On the other hand, if we fix the number of quadrature points (i.e., poles) and refine via filtered subspace iteration, the residuals of all target eigenpairs converge geometrically to order u.

The aim of this paper is to explain Figures 1.1 and 1.2. Our analysis generalizes that of Peters and Wilkinson [10], who demonstrated that single-vector inverse iterations converge rapidly when the shift is very close to an eigenvalue, to the substantially more complex case of subspace iteration and rational filters with multiple poles. We first examine how rational filtered subspace iteration disarms dangerous eigenvalues after the first iteration. When A has a complete set of orthonormal eigenvectors, orthogonal bases for the search subspace play a special role and "twice is enough" to recover full precision in the computed iterates (see sections 3, 4, and 5).[1] In the nonnormal case, iterating on approximate eigenvectors (obtained from a Rayleigh–Ritz step, for instance) is the key to overcoming round-off errors incurred by the dangerous eigenvalue, while iterations based on orthogonal bases (such as approximate Schur vectors) suffer stagnation in the remaining target eigenpairs (see section 6). For nonnormal matrices, finite-precision effects may persist after two iterations.

To obtain full precision in the remaining target eigenpairs for Arnoldi and related Krylov schemes, the prevailing consensus is to alter the rational filter by moving or removing the offending poles [1, 15]. Unfortunately, this usually means settling for a less efficient filter or starting over with a new filter. Informed by our analysis of subspace iteration and its immunity to dangerous eigenvalues, we offer simple restart strategies that fix stagnation in shift-and-invert Arnoldi (see section 7).

Our analysis is focused on a matrix A with a single dangerous eigenvalue, λ, located at a distance $d \ll 1$ from a pole z_{j_*} (for some $1 \le j_* \le \ell$) of the filter in (1.2). To reveal the precise influence of the dangerous eigenvalue, we frame our discussion in the asymptotic limit

$$d = |z_{j_*} - \lambda| \to 0. \tag{1.3}$$

[1] Aspects of our analysis are similar to Parlett and Kahan's "twice-is-enough" algorithm and analysis for Gram–Schmidt reorthogonalization [9, sect. 6.9].

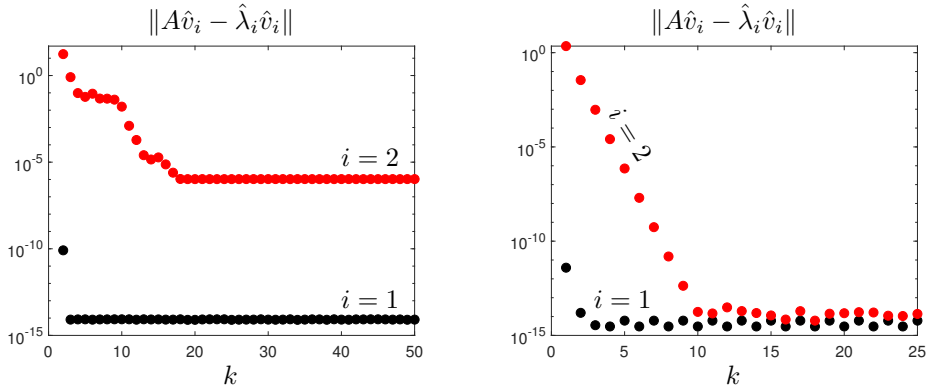

FIG. 1.1. *The residuals for two approximate eigenpairs of a real-symmetric* 100×100 *matrix at iterations* $k = 2, \ldots, 50$ *of Arnoldi (left) and iterations* $k = 1, \ldots, 25$ *of subspace iteration (right), both with shift-and-invert enhancement. The approximate eigenpairs correspond to a dangerous eigenvalue (black) with* $|z - \lambda_1| = 10^{-12}$ *and a second target eigenvalue (red) with* $|z - \lambda_2| \approx 0.1$. *(See online version for color.)*

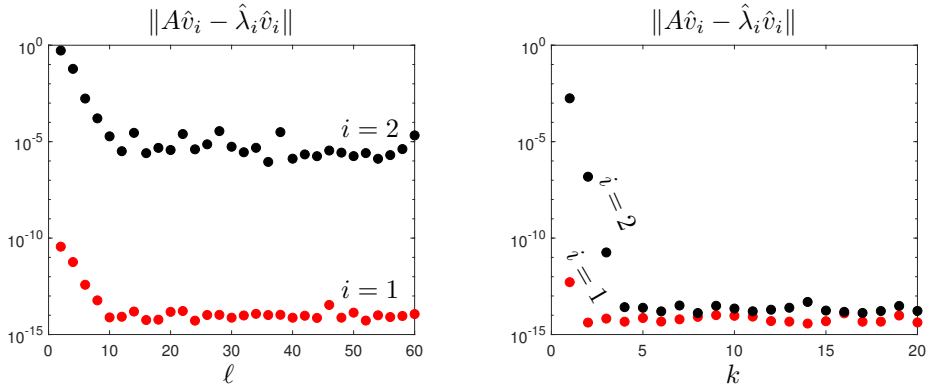

FIG. 1.2. *The left panel displays the residuals for two approximate eigenpairs of a* 100×100 *real-symmetric matrix computed with the contour integral eigensolver described in* [17], *as the quadrature rule approximating the contour integral is refined. One of the target eigenvalues (*$i = 1$, *red) is a distance of* 10^{-10} *from the contour. The right panel displays the residuals for the two approximate eigenpairs when refined via iteration rather than quadrature rule. The quadrature rule used corresponds to a rational filter in* (1.2) *with* $\ell = 8$. *(See online version for color.)*

However, we always provide concrete bounds and give leading order estimates to elucidate the role of salient parameters, e.g., those related to the rational filter or nonnormality of A. We consider the implications of our results for other natural configurations, such as multiple eigenvalues clustered at a pole, in section 8.

Throughout the paper, $\|\cdot\|$ denotes the spectral norm of a matrix (Euclidean norm for vectors) and A denotes an $n \times n$ diagonalizable matrix with eigenvalues and eigenvectors satisfying $Av_i = \lambda_i v_i$ for $1 \leq i \leq n$. Except in section 6, we assume that A has a complete orthonormal set of eigenvectors (i.e., A is normal), in which case it is convenient to write the eigendecomposition of A in the form

$$A = V_1 \Lambda_1 V_1^* + V_2 \Lambda_2 V_2^*. \tag{1.4}$$

Here, $\Lambda_1 = \mathrm{diag}(\lambda_1, \ldots, \lambda_m)$ contains a set of target eigenvalues that we wish to compute, and $\Lambda_2 = \mathrm{diag}(\lambda_{m+1}, \ldots, \lambda_n)$ contains the remaining unwanted eigenvalues (usually, $m \ll n$). We denote the target eigenspace by $\mathcal{V} = \mathrm{span}(V_1)$, the full eigen-

value matrix of A by $\Lambda = \mathrm{diag}(\Lambda_1, \Lambda_2)$, and the eigenvector matrix by $V = [V_1\ V_2]$.

For simplicity, we always assume that $r(\Lambda)$ is invertible and that there is a nonzero spectral gap between $r(\Lambda_1)$ and $r(\Lambda_2)$, and we index the eigenvalues in order of decreasing modulus under the filter so that

$$|r(\lambda_1)| \geq \cdots \geq |r(\lambda_m)| > |r(\lambda_{m+1})| \geq \cdots \geq |r(\lambda_n)|. \tag{1.5}$$

Here, $r(\lambda) = \sum_{j=1}^{\ell} \omega_j (z_j - \lambda)^{-1}$ is the scalar form of the filter in (1.2).[2] Under the ordering in (1.5), the dangerous eigenvalue is λ_1. Without loss of generality, we assume that the weight w_j associated with a pole near the dangerous eigenvalue λ_1 is equal to one (by scaling $r(\cdot)$ if necessary). This simplifies the analysis and usually implies that the other weights w_i are also modest in size. Finally, we tacitly assume $\|A\| = \mathcal{O}(1)$ in informal discussions; the formal theorems and statements hold without this assumption. The only place where $\|A\|$ appears in the analysis is section 5.

2. Subspace iteration with rational filters. Given an $n \times m$ matrix Q_0 with orthonormal columns, the simplest practical form of subspace iteration with a rational filter, as in (1.2), computes the iterates

$$X_k = r(A)Q_{k-1}, \qquad Q_k = \mathrm{qf}(X_k). \tag{2.1}$$

Here, $\mathrm{qf}(X_k)$ denotes the orthogonal factor from a QR decomposition of X_k. The eigenvalues of $Q_k^* A Q_k$ provide approximations to the target eigenvalues, and approximate eigenvectors are given by $Q_k x_i$ for each eigenvector, x_i, of the small $m \times m$ matrix $Q_k^* A Q_k$. These approximations to the target eigenpairs are called Ritz pairs.

Intuitively, the Ritz pairs extracted with the basis Q_k are usually good approximations to the target eigenpairs when there are good approximations to $v_1, \ldots, v_m$ in $\mathcal{S}_k = \mathrm{span}(Q_k)$. Here, the rational filter in (2.1) fills two complementary roles. First, the filter should guide the iterates toward the target eigenspace by mapping the target eigenvalues of A to the dominant eigenvalues of $r(A)$ (that is, the eigenvalues with the largest modulus $|r(\lambda_i)|$). Second, the filter should accelerate the convergence of the Ritz pairs by ensuring that $|r(\lambda_{m+1})/r(\lambda_m)| \ll 1$. These criteria follow from a standard one-step refinement bound for subspace iteration [12, Thm. 5.2].

THEOREM 2.1. *Let normal $A \in \mathbb{C}^{n\times n}$ and $r : \Lambda \to \mathbb{C}$ satisfy (1.4) and (1.5), respectively, and let $\mathcal{S}_k = \mathrm{span}(Q_k)$ in (2.1) for $k \geq 0$. If $V_1^* Q_0$ has full rank, then for each $v_i \in \mathcal{V}$ there are vectors $s_i^{(k)} \in \mathcal{S}_k$ such that*

$$\|s_i^{(k)} - v_i\| \leq \left|\frac{r(\lambda_{m+1})}{r(\lambda_i)}\right| \|s_i^{(k-1)} - v_i\| \leq \left|\frac{r(\lambda_{m+1})}{r(\lambda_i)}\right|^k \|s_i^{(0)} - v_i\|. \tag{2.2}$$

Moreover, each $P_{\mathcal{V}} s_i^{(k)} = v_i$, where $P_{\mathcal{V}} = V_1 V_1^*$ is the spectral projector onto $\mathcal{V}$.

Theorem 2.1 implies that there are approximations in $\mathcal{S}_k$ that converge geometrically to the ith target eigenvector with rate $|r(\lambda_{m+1})|/|r(\lambda_i)|$.[3] Consequently, if the filter is very small on the unwanted eigenvalues relative to its magnitude on the target eigenvalues, then we expect the Ritz pairs to converge rapidly. The appeal of rational filters in the modern computing era is that filters of modest degree $\ell \leq 20$

[2] We refer to the scalar function $r(z)$ and its matrix companion $r(A)$ with the same symbol. We always include the argument when it is necessary to clarify which one we mean.

[3] If A does not possess orthogonal eigenvectors, this rate is only asymptotic as $k \to \infty$ due to the phenomenon of transient growth in matrix powers of nonnormal matrices [21, Chap. 16].

often achieve $|r(\lambda_{m+1})|/|r(\lambda_m)| \approx u$. In a typical parallel computing environment, the individual shifted inverses in (1.2) are easily applied in parallel, meaning that the target eigenpairs can be computed to machine precision at the equivalent (serial) cost of solving a shifted linear system. However, a higher degree rational filter and multiple iterations may be required when many eigenvalues are clustered near the target group. Additionally, clustered eigenvalues may lead to ill-conditioned eigenvectors and loss of orthogonality in the Ritz pairs. When eigenvalues are clustered and more poles are employed in the rational filter, one may also encounter dangerous eigenvalues.

In practice, there are many modifications one can make to (2.1) to improve convergence, enhance stability, or increase computational efficiency. Nevertheless, when A is normal, (2.1) is enough to capture both the dangers and the self-correcting effects of eigenvalues that are close to the poles in (1.2).

2.1. Principal angles between subspaces. The principal angles between the subspaces $\mathcal{S}_k$ and $\mathcal{V}$ provide a natural framework with which to characterize the refinement of the iterates in (2.1). Generalizing the notion of an angle between two vectors, the principal angles tell us how close $\mathcal{S}_k$ and $\mathcal{V}$ are in a geometric sense [2].

DEFINITION 2.2. *Let $\mathcal{X}$ and $\mathcal{Y}$ be two m-dimensional subspaces with orthonormal bases X and Y, respectively, and let $\sigma_i(Y^*X)$ denote the ith singular value of Y^*X. The principal angles between $\mathcal{X}$ and $\mathcal{Y}$ are the acute angles $\theta_1(\mathcal{X},\mathcal{Y}) \geq \cdots \geq \theta_m(\mathcal{X},\mathcal{Y})$ satisfying*

$$\cos\theta_i(\mathcal{X},\mathcal{Y}) = \sigma_{m+1-i}(Y^*X), \qquad i = 1,\dots,m. \tag{2.3}$$

The sine of the largest principal angle, given by $\sin\theta_1(\mathcal{X},\mathcal{Y}) = \|(I - P_{\mathcal{Y}})X\|$, defines a metric on the set of m-dimensional subspaces. However, the tangents of the principal angles, which are the singular values of the matrix [23]

$$T(X,Y) = (I - P_{\mathcal{Y}})X(Y^*X)^+, \tag{2.4}$$

are better equipped to describe the behavior of the iterates in (2.1). In (2.4), $(Y^*X)^+$ denotes the Moore–Penrose pseudoinverse of Y^*X, and, crucially, X need not be orthonormal, that is, $T(X,Y) = T(Q,Y)$ where $X = QR$ is the thin QR factorization. A subspace analogue of Theorem 2.1, based on the largest principal angle between $\mathcal{S}_k$ and $\mathcal{V}$, is easy to derive with (2.4) (cf. [9, Thm. 14.4.1]).

THEOREM 2.3. *Let normal $A \in \mathbb{C}^{n\times n}$ and $r : \Lambda \to \mathbb{C}$ satisfy* (1.4) *and* (1.5)*, respectively, and let $\mathcal{S}_k = \operatorname{span}(Q_k)$ in* (2.1)*. If $\cos\theta_1(\mathcal{S}_0,\mathcal{V}) > 0$, then*

$$\tan\theta_1(\mathcal{S}_k,\mathcal{V}) \leq \left|\frac{r(\lambda_{m+1})}{r(\lambda_m)}\right| \tan\theta_1(\mathcal{S}_{k-1},\mathcal{V}) \leq \left|\frac{r(\lambda_{m+1})}{r(\lambda_m)}\right|^k \tan\theta_1(\mathcal{S}_0,\mathcal{V}). \tag{2.5}$$

Proof. We prove the first inequality with a direct calculation using (2.4); the second follows immediately by induction and the fact that $\cos\theta > 0$ when $\tan\theta < \infty$. We compute that $(I - P_{\mathcal{V}})X_k = V_2 r(\Lambda_2) V_2^* Q_{k-1}$ and that $V_1^* X_k = r(\Lambda_1) V_1^* Q_{k-1}$. Using the induction hypothesis that $\cos\theta_1(\mathcal{S}_{k-1},\mathcal{V}) > 0$, which implies $V_1^* Q_{k-1}$ is invertible, we obtain

$$(I - P_{\mathcal{V}})X_k(V_1^*X_k)^+ = V_2 r(\Lambda_2) V_2^* Q_{k-1}(V_1^*Q_{k-1})^{-1} r(\Lambda_1)^{-1}. \tag{2.6}$$

The theorem follows by taking norms and noting that $\|V_2^*Q_{k-1}(V_1^*Q_{k-1})^{-1}\| = \tan\theta_1(\mathcal{S}_{k-1},\mathcal{V})$, $\|r(\Lambda_2)\| = |r(\lambda_{m+1})|$, and $\|r(\Lambda_1)^{-1}\| = |r(\lambda_m)|^{-1}$. ∎

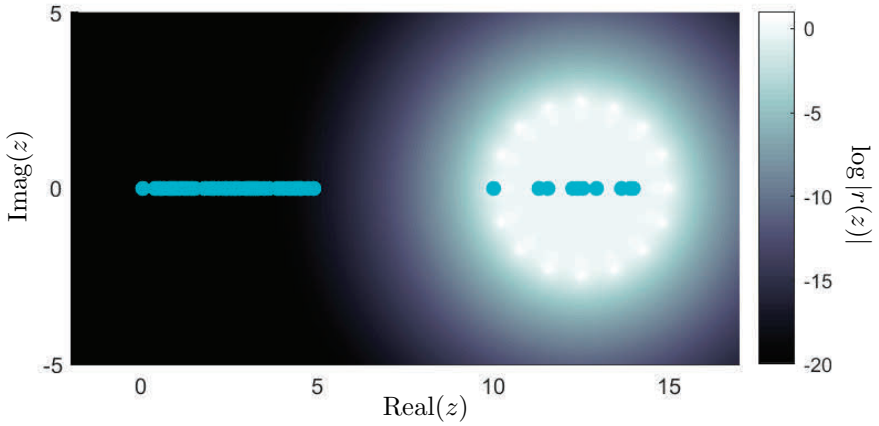

FIG. 3.1. *The eigenvalues of a* 100×100 *real-symmetric matrix overlaid on a complex color plot of the magnitude of a rational approximation to the characteristic function on* $[10, 15]$. *A dangerous eigenvalue is located at distance* $d = 10^{-10}$ *from the pole at* $z = 10$.

We note that Theorem 2.1 is recovered from (2.6) by postmultiplying each side by the unit vector e_i and setting $s_j = X_k(V_1^* X_k)^+ e_i$ for $j = k-1, k$.

The tangents (and sines) of the principal angles play an important role in the perturbation theory of eigenpairs, and, consequently, the bounds in Theorem 2.3 are useful when determining the accuracy in the computed Ritz pairs [12, 15, 16]. For our purposes, Theorem 2.3 and its proof are useful tools when analyzing subspace iterations subject to perturbations (see section 5) because $\tan\theta_1(\mathcal{S}_k, \mathcal{V})$ is computed directly from the iterate X_k.

3. Dangerous eigenvalues. When an eigenvalue of A is very close to a pole of the rational filter in (1.2), $r(A)$ disproportionately amplifies components in the direction of the associated eigenvector. To see this, note that $r(\lambda_1) = (z_{j_*} - \lambda_1)^{-1} + \sum_{j=1, j\neq j_*}^{\ell} w_j (z_j - \lambda_1)^{-1}$ (recall that $w_{j_*} = 1$ by assumption) and that $r(\lambda_i) = \mathcal{O}(1)$ for $i \geq 2$ since there is only one dangerous eigenvalue. Then, given any vector $x \in \mathbb{C}^n$,

$$r(A)x = \sum_{i=1}^{n} r(\lambda_i) v_i v_i^* x = \frac{v_1^* x}{d e^{i\theta}} v_1 + \mathcal{O}(1) \quad \text{as} \quad d \to 0. \tag{3.1}$$

(It is convenient to write the complex-valued difference between λ_1 and the nearest pole z_{j_*} in the polar notation $z_{j_*} - \lambda_1 = d e^{i\theta}$, with argument $0 \leq \theta < 2\pi$.) This amplification is precisely the reason that shift-and-invert power iterations are so effective when the shift is close to the target eigenvalue. If we apply $r(A)$ to a random vector with unit norm and normalize, the result approximates v_1 with relative accuracy $\mathcal{O}(d)$, under the generic assumption that the random vector is not nearly orthogonal to v_1. Similarly, when $r(A)$ is applied to a random orthonormal matrix Q_0, $\operatorname{span}(r(A)Q_0)$ contains good approximations to v_1 when $\|v_1^* Q_0\|$ is not too small.

However, the amplifying effect of a dangerous eigenvalue may cause issues when computing the iterates in (2.1) in floating-point arithmetic. Figure 3.1 shows the eigenvalues of a 100×100 real-symmetric matrix plotted in the complex plane over the magnitude (indicated by color) of a rational filter targeting the interval $[10, 15]$. The matrix has a large cluster of eigenvalues in the interval $[0, 5]$, where the filter has decayed to less than unit round-off and a small set of eigenvalues in the target region,

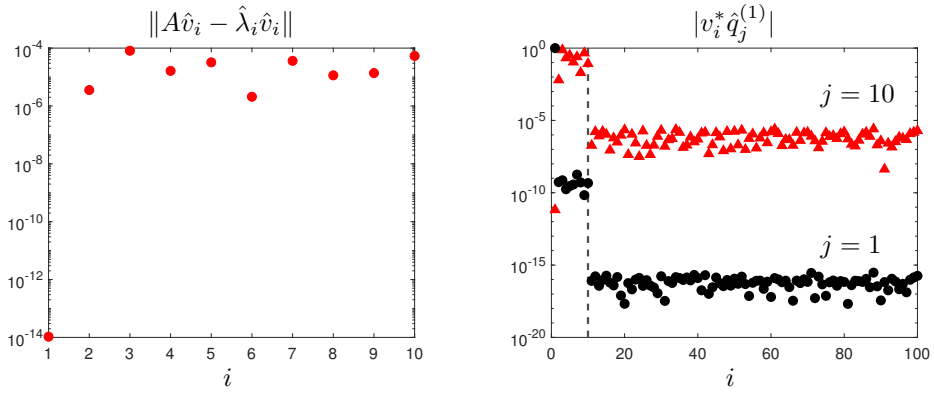

FIG. 3.2. *Plotted on the left are the residuals of* 10 *target eigenpairs of a* 100×100 *real-symmetric matrix after one iteration of subspace iteration with the rational filter in Figure* 3.1. *On the right are the eigenvector coordinates of the* 1*st (black circles) and* 10*th (red triangles) columns of* $\hat{Q}_1$. *The dangerous component and the remaining target components dominate in columns* 1 *and* 10*, respectively. The unwanted eigenvector components (right of dashed line) are filtered out almost entirely to order* u *in the* 1*st column but are orders of magnitude larger in the* 10*th column, with magnitude near* u/d.

where the filter has magnitude close to 1. One eigenvalue of the matrix is very close to the pole at $z = 10$, separated by a distance of 10^{-10}. By Theorem 2.1, we expect that (in exact arithmetic) all of the eigenvalues in the target region are resolved to accuracy on the order of u after one iteration. However, Figure 3.2 (left) shows that only the dangerous eigenpair has been computed accurately. The residuals of the remaining target eigenpairs are on the order of 10^{-5}, that is, roughly u/d.

The large residuals are best understood by looking at the computed orthonormal basis $\hat{Q}_1$ (we denote computed quantities with a hat throughout, so $\hat{Q}_1$ is the computed approximant to Q_1) in the eigenvector coordinates in Figure 3.2. The first column of $\hat{Q}_1$ (circular markers) appears as expected: the dangerous eigenvector dominates, and the unwanted components are near the unit round-off in magnitude. However, the magnitude of the unwanted components is much larger, on the order of u/d, in the remaining columns $\hat{q}_2^{(1)}, \ldots, \hat{q}_m^{(1)}$. The 10th column (triangular markers in Figure 3.2) is representative of this observation. Although the quality of the filter means that the unwanted components should be on the order of u in the columns of $\hat{Q}_1$, they are polluted with noise on the order of u/d in all but the first column. Consequently, the accuracy in the remaining Ritz pairs computed from $\hat{Q}_1$ is also degraded to u/d.

There are two potential sources of error degrading the accuracy in $\hat{Q}_1$. The first is the most obvious: round-off errors are amplified when solving the ill-conditioned linear system associated with the dangerous eigenvalue. The second source is more subtle: the overwhelming dominance of the dangerous eigenvector in each column of X_1 leads to an ill-conditioned basis for $\mathcal{S}_1$. Remarkably, the heart of the story in Figures 1.1 and 1.2 is contained in the latter, subtler effect. We address both points in section 5, where we study the convergence and stability of the iteration in (2.1) when computed in floating-point arithmetic. For now, we focus on the influence of ill-conditioning in the iterates $X_1, X_2, \ldots$, noting that round-off errors in the computed iterates have little effect on their condition number (see section 5 for a full explanation).

3.1. Accuracy of the computed orthonormal basis. When a basis $X \in \mathbb{C}^{n\times m}$ is ill-conditioned, small perturbations to the columns can have a large effect on their span. This is reflected in the sensitivity of the orthogonal factor in the QR

factorization, $Q = \mathrm{qf}(X)$. If, for some small $\epsilon > 0$, X is perturbed by ΔX with $\|\Delta X\| \le \epsilon \|X\|$, then there is a ΔQ such that $Q + \Delta Q = \mathrm{qf}(X + \Delta X)$ and [5, p. 382]

$$\|\Delta Q\| \le c_m \kappa(X) \|\Delta X\| / \|X\|. \tag{3.2}$$

Here, c_m is a modest constant depending only on the dimension m, and $\kappa(\cdot)$ denotes the 2-norm condition number of a rectangular matrix. Equation (3.2) tells us that when X is highly ill-conditioned, the QR factorization may be extremely sensitive to perturbations. Consequently, when we compute an orthonormal basis $\hat{Q}$ in floating-point arithmetic, we are not guaranteed accuracy much better than $\|\hat{Q} - Q\| \le c_m \kappa(X) u$.

There is an important caveat to (3.2): the basis Q and its sensitivity are unaffected when the columns of X are rescaled [5, sect. 18.8]. A near-optimal scaling equilibrates the columns to have unit norm, minimizing $\kappa(XT)$ to within a factor of $\sqrt{m}$ over all diagonal matrices T [5, sect. 7.3]. In other words, the bound in (3.2) is descriptive when the columns of X do not vary significantly in magnitude. While large column norms and small angles between columns both contribute to $\kappa(X)$, only the latter affect ΔQ. Column scaling plays an important role in the analysis of the second iteration in subsection 4.1. However, it has little effect on the condition number of $\kappa(X_1)$ because the columns of X_1 are all of order $1/d$, as we now explain.

Because the rational filter amplifies the v_1 component in each column of Q_0 by $1/d$ in (2.1), X_1 is usually extremely ill-conditioned. Intuitively, $\kappa(X_1)$ cannot be much worse than $|r(\lambda_1)|/|r(\lambda_m)|$ and not much better than $|r(\lambda_1)|/|r(\lambda_2)|$ because v_1 is present in each column with magnitude near $|r(\lambda_1)|$, while the rest of the target eigenpairs are present with magnitude at least $|r(\lambda_m)|$ and no greater than $|r(\lambda_2)|$. Proposition 3.1 makes this intuition precise in the form of an upper bound and asymptotic lower bound. The implication is that the error in the computed orthonormal basis $\hat{Q}_1$ is on the order of u/d as long as the columns of Q_0 are not orthogonal to the dangerous eigenvector, as we observed in Figure 3.2.

We use the shorthand notation $f(x) \lesssim g(x)$ to denote the asymptotic relation

$$f(x) \le g(x)(1 + o(1)) \quad \text{as} \quad x \to 0. \tag{3.3}$$

This is sharper than $f(x) = \mathcal{O}(g(x))$ but weaker than $f(x) \sim g(x)$ [6, Defs. 1.1–1.2].[4]

PROPOSITION 3.1. *Let normal $A \in \mathbb{C}^{n \times n}$ and $r : \Lambda \to \mathbb{C}$ satisfy* (1.4) *and* (1.5)*, respectively, and, given orthonormal $Q_0 \in \mathbb{C}^{n \times m}$, let $X_1 = r(A)Q_0$. If $V_1^* Q_0$ has full rank, then the condition number of X_1 satisfies*

$$\frac{\|v_1^* Q_0\|}{d|r(\lambda_2)|} \lesssim \kappa(X_1) \le \left|\frac{r(\lambda_1)}{r(\lambda_m)}\right| \|(V_1^* Q_0)^{-1}\| \quad \textit{as} \quad d \to 0. \tag{3.4}$$

Proof. The condition number of X_1 may be written as $\kappa(X_1) = \sigma_1(X_1)/\sigma_m(X_1)$, where $\sigma_1(X_1) \ge \cdots \ge \sigma_m(X_1)$ are the singular values of X_1. To bound $\sigma_1(X_1)$ above, we substitute the spectral decomposition $r(A) = V r(\Lambda) V^*$ into the definition of X_1 and estimate $\sigma_1(X_1) \le |r(\Lambda_1)| \|V^* Q_0\| \le |r(\lambda_1)|$. To bound $\sigma_m(X_1)$ below, we use the spectral decomposition in (1.4) to write

$$X_1 = \begin{bmatrix} V_1 & V_2 \end{bmatrix} \begin{bmatrix} M_1 \\ M_2 \end{bmatrix}, \tag{3.5}$$

[4] This definition of $f \lesssim g$ is sharper than its common usage in the analysis of partial differential equations, where it means $f \le Cg$ for some constant $C > 0$ [18, p. xiv].

where $M_1 = r(\Lambda_1)V_1^*Q_0$ and $M_2 = r(\Lambda_2)V_2^*Q_0$. Because V is unitary, the singular values of X_1 are precisely those of $\left[\begin{smallmatrix} M_1 \\ M_2 \end{smallmatrix}\right]$. Furthermore, $\sigma_m(X_1) \geq \sigma_m(M_1)$ since adding rows can only increase the singular values of a matrix. Finally, since $\sigma_m(M_1) = \|M_1^{-1}\|^{-1}$, we have that $\kappa(X_1) \leq \sigma_1(X_1)\|M_1^{-1}\|$. We estimate that

$$\|M_1^{-1}\| = \|(V_1^*Q_0)^{-1}r(\Lambda_1)^{-1}\| \leq \|(V_1^*Q_0)^{-1}\||r(\lambda_m)|^{-1}.$$

Collecting the bounds on $\sigma_1(X_1)$ and $\|M_1^{-1}\|$ establishes the upper bound in (3.4).

To establish the asymptotic lower bound, we apply (3.1) to $r(A)Q_0$, obtaining

$$X_1 = \sum_{i=1}^{n} r(\lambda_i)v_i v_i^* Q_0 = v_1 \frac{v_1^* Q_0}{de^{i\theta}} + \mathcal{O}(1) \quad \text{as} \quad d \to 0. \tag{3.6}$$

Taking norms provides the asymptotic lower bound on $\sigma_1(X_1)$. To obtain a lower bound on $\sigma_m(X_1)^{-1}$, we can bound $\sigma_m(X_1)$ from above with an interlacing property for singular values of matrices subject to rank one perturbations. We rewrite (3.6) as

$$X_1 = r(\lambda_1)v_1 v_1^* Q_0 + V\,\mathrm{diag}(0, r(\lambda_2), \ldots, r(\lambda_n))V^*Q_0 = N_1 + N_2.$$

Now, $\sigma_2(N_1) = 0$ and $\sigma_1(N_2) \leq |r(\lambda_2)|$, so by interlacing [19], we obtain the estimate

$$\sigma_2(X_1) \leq \sigma_1(N_2) + \sigma_2(N_1) \leq |r(\lambda_2)|.$$

As $\sigma_m(X_1) \leq \sigma_2(X_1)$ implies that $1/\sigma_m(X_1) \geq |r(\lambda_2)|^{-1}$, collecting lower bounds concludes the proof of (3.4). ∎

The factor $\|(V_1^*Q_0)^{-1}\|$ in Proposition 3.1 appears naturally in connection with subspace iteration, and we will encounter it again in section 5. It is precisely the reciprocal of $\cos\theta_1(\mathcal{S}_0, \mathcal{V})$ (see Definition 2.2), approaching unity when $\mathcal{V}$ and $\mathcal{S}_0$ are nearby, and blowing up quadratically when they are made orthogonal. In Proposition 3.1 it indicates that X_1 may suffer additional ill-conditioning if the initial subspace $\mathcal{S}_0$ is accidentally chosen to be too near orthogonal to $\mathcal{V}$.[5]

4. Twice is enough. In Proposition 3.1, the asymptotic lower bound in (3.4) plummets if the columns of Q_0 are taken nearly orthogonal to v_1, the dangerous eigenvector. This is because the rational filter has nothing to amplify when v_1 is absent in the columns of Q_0. If v_1 is present with magnitude no greater than $\mathcal{O}(d)$ in Q_0, then the columns of X_1 are not strongly aligned along any single eigenvector, and the conditioning of X_1 is likely to improve. Crucially, this intuition holds even if v_1 dominates in one column but not the others. The main point is that the columns of X_1 are no longer necessarily close to a linearly dependent set.

Let us return to the example of Figure 3.1. If we print out the residual norms of the target eigenpairs after the second iteration of (2.1), we see remarkable improvement:

```
6.7997e-15   2.5942e-14   2.2680e-13   4.3433e-14   9.1978e-14
1.3716e-14   9.7045e-14   3.4121e-14   1.4594e-13   4.0235e-14
```

Now all the target pairs have been resolved to within 13 or 14 digits of accuracy, in contrast to Figure 3.2 (left). If we examine the computed orthonormal basis used

[5] When Q_0 is selected so that its entries are independent and identically distributed Gaussian random variables, $\|(V_1^*Q_0)^{-1}\|$ is roughly $\sqrt{m}$ in expectation but can be an order of magnitude or so larger with nontrivial probability. A powerful workaround is to work with a slightly larger subspace and take m larger than the number of target eigenvalues; this dramatically reduces the probability of large $\|(V_1^*Q_0)^{-1}\|$ [3].

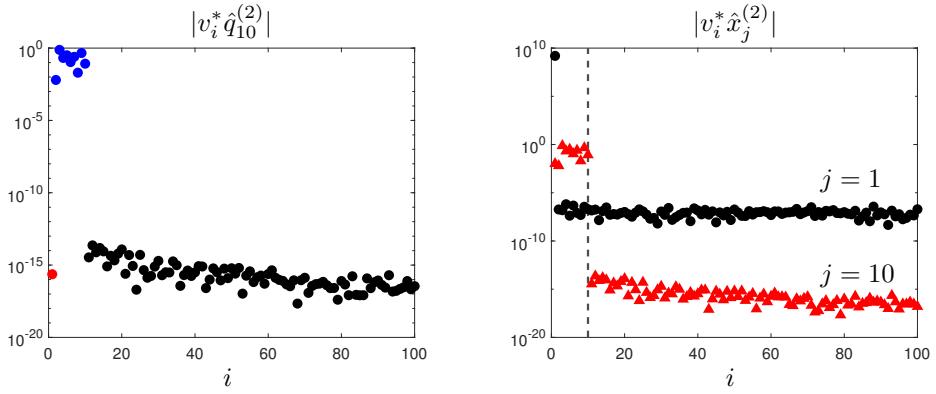

FIG. 4.1. *The structure of the iterates $\hat{X}_2$ and $\hat{Q}_2$ after the second iteration of subspace iteration with a rational filter. On the left are the eigenvector coordinates of the 10th column of the computed orthonormal basis color-coded for the dangerous component (red), remaining target components (blue), and unwanted components (black). On the right are the eigenvector coordinates of the 1st (black circles) and 10th (red triangles) columns of the computed basis $\hat{X}_2$, with a dashed line dividing target and unwanted eigenvector coordinates. (See online version for color.)*

to extract the Ritz pairs, we observe that the noise in the direction of the unwanted eigenvectors has also been reduced to the order of u, compared with u/d in the first iteration. Figure 4.1 (left) illustrates the composition of the 10th column of $\hat{Q}_2$, which is representative of the last $m-1$ columns.

The reason for the restored accuracy in the computed orthonormal basis is that, unlike X_1, the basis X_2 has an even blend of the target eigenvector directions in all but the first of its columns. Figure 4.1 (right) displays the magnitude of the eigenvector coordinates for the first (circular markers) and last (triangular markers) columns of the computed basis, $\hat{X}_2$. In the first column, the dangerous direction is effectively the only direction present. In contrast, the last column of $\hat{X}_2$ contains order one components in each target direction, with the unwanted directions completely filtered out. The remaining columns of $\hat{X}_2$ are similar in composition to the last. Without v_1 dominating in every column, we can accurately extract an orthonormal basis.

The clue to the stark difference in the composition of $\hat{X}_1$ and $\hat{X}_2$ is contained in Figure 3.2. We see that the first column of $\hat{Q}_1$ is dominated by the dangerous eigenvector, up to the 9th or 10th digit. Consequently, the remaining columns of $\hat{Q}_1$ are nearly orthogonal to v_1. We observe this in Figure 3.2 (right), where $v_1^* q_{10}^{(1)} \approx 10^{-11}$. When the rational filter is applied to $\hat{Q}_1$ in the second iteration, the amplification of v_1 restores an even blend of the target eigenvectors in the last $m-1$ columns of $\hat{X}_2$, rather than boosting v_1 above the others.

4.1. A well-conditioned basis. Motivated by the preceding discussion, we now examine the columns of X_2, when $q_1^{(1)}$ is a good approximation to v_1, and consequently all but one of the columns of Q_1 are deficient in the dangerous direction. Unlike X_1, the columns of X_2 vary between $\mathcal{O}(1)$ and $\mathcal{O}(1/d)$ in magnitude. Therefore, we estimate the condition number of X_2 after rescaling all columns to $\mathcal{O}(1)$ (see the discussion following (3.2)). We then briefly explain the structure of the eigenvector coordinates for the computed orthonormal basis $\hat{Q}_1$ observed in Figure 3.2.

To investigate how a weak presence of v_1 in columns of Q_1 improves the condi-

tioning of X_2, we break the target eigenvector coordinates of Q_1 into blocks as

$$V_1^* Q_1 = \begin{bmatrix} v_1^* q_1^{(1)} & v_1^* \tilde{Q}_1 \\ \tilde{V}_1^* q_1^{(1)} & \tilde{V}_1^* \tilde{Q}_1 \end{bmatrix} = \begin{bmatrix} a & b \\ c & D \end{bmatrix}. \tag{4.1}$$

Here, we use $\tilde{V}_1$ and $\tilde{Q}_1$ to denote the $n \times (m-1)$ matrices formed by removing the first columns of V_1 and Q_1, respectively. When $q_1^{(1)}$ closely approximates v_1, $\|c\|$ is small due to the orthogonality of the eigenvectors. Moreover, $\|b\|$ is also small because the columns of $\tilde{Q}_1$ are nearly orthogonal to v_1. Let us suppose that $q_1^{(1)} = v_1 + \mathcal{O}(d)$, so that b and c are $\mathcal{O}(d)$ (we explain why this holds, following Theorem 4.1).

After applying the rational filter to Q_1, the eigenvector coordinates $V_1^* X_2$ inherit a natural block structure from (4.1). Letting $\tilde{\Lambda}_1 = \operatorname{diag}(\lambda_2, \ldots, \lambda_m)$, we have that

$$V_1^* X_2 = r(\Lambda_1) V_1^* Q_1 = \begin{bmatrix} r(\lambda_1)a & r(\lambda_1)b \\ r(\tilde{\Lambda}_1)c & r(\tilde{\Lambda}_1)D \end{bmatrix}. \tag{4.2}$$

While the bottom left block remains small, with norm no greater than $|r(\lambda_2)|\|c\| = \mathcal{O}(d)$, the entire first row is amplified by $|r(\lambda_1)|$ so that $\|r(\lambda_1)b\| \approx \|b\|/d$.

Before estimating the condition number of X_2, we scale the columns with the $m \times m$ diagonal matrix

$$T = \operatorname{diag}(r(\lambda_1)^{-1}, 1, \ldots, 1). \tag{4.3}$$

This diagonal scaling does not alter $\operatorname{span}(X_2)$ or the sensitivity of the orthonormal basis, $Q_2 = \operatorname{qf}(X_2)$ [5, sect. 18.8], but it forces the columns of $X_2 T$ to have approximately equal norms. This conveniently puts the diagonal blocks in (4.2) on equal footing and ensures that $\sigma_1(X_2 T) = \mathcal{O}(1)$, so that any ill-conditioning due to the dangerous eigenvalue is captured in the smallest singular value of $X_2 T$. This allows us to focus on computing a lower bound for $\sigma_m(X_2 T)$ or, equivalently, an upper bound for $1/\sigma_m(X_2 T)$. Just as in the proof of Proposition 3.1, it suffices to bound $\|(V_1^* X_2 T)^{-1}\|$ from above (assuming as usual that $V_1^* Q_1$, and therefore $V_1^* X_2$, has full rank).

With all of the ingredients in place, the estimate is fairly straightforward. After the column scaling, $V_1^* X_2 T$ is approximately block upper triangular. We have that

$$V_1^* X_2 T = \begin{bmatrix} r(\lambda_1)a & r(\lambda_1)b \\ r(\tilde{\Lambda}_1)c & r(\tilde{\Lambda}_1)D \end{bmatrix} T = \begin{bmatrix} a & b/(de^{i\theta}) \\ & r(\tilde{\Lambda}_1)D \end{bmatrix} + \mathcal{O}(d). \tag{4.4}$$

We can apply the formula for 2×2 block upper triangular matrix inversion and the fact that matrix inversion is locally Lipschitz continuous to compute (for d sufficiently small)

$$(V_1^* X_2 T)^{-1} = \begin{bmatrix} a^{-1} & -a^{-1} D^{-1} r(\tilde{\Lambda}_1)^{-1} b/(de^{i\theta}) \\ & D^{-1} r(\tilde{\Lambda}_1)^{-1} \end{bmatrix} (I + \mathcal{O}(d)). \tag{4.5}$$

The norm of the block upper triangular matrix in (4.5) is bounded by the sum of the norms of the blocks, so we conclude that $1/\sigma_m(X_2 T) \leq \|(V_1^* X_2 T)^{-1}\| = \mathcal{O}(1)$ when $\|b\| = \mathcal{O}(d)$. Estimating the norms of these blocks individually and combining with an estimate for $\sigma_1(X_2 T)$ leads to the following upper bound on $\kappa(X_2 T)$.

THEOREM 4.1 (twice-is-enough). *Let normal $A \in \mathbb{C}^{n \times n}$ and $r : \Lambda \to \mathbb{C}$ satisfy* (1.4) *and* (1.5)*, respectively, and, given orthonormal $Q_1 \in \mathbb{C}^{n \times m}$, let $X_2 =$*

$r(A)Q_1$. Let b and D denote the blocks of $V_1^*Q_1$ in (4.1). If D is invertible and the first column of Q_1 satisfies $\|q_1^{(1)} - v_1\| = \mathcal{O}(d)$, then $\|b\| = \mathcal{O}(d)$ and

$$\kappa(X_2T) \leq M\left(\left(\frac{\|b\|}{d}+1\right)\frac{\|D^{-1}\|}{|r(\lambda_m)|}+1\right)+\mathcal{O}(d) \quad \text{as} \quad d\to 0. \tag{4.6}$$

Here, $T = \operatorname{diag}(r(\lambda_1)^{-1}, 1, \ldots, 1) \in \mathbb{C}^{m\times m}$ and $M = \|b\|/d + \max\{1, |r(\lambda_2)|\}$.

Proof. First, the hypothesis $\|q_1^{(1)} - v_1\| = \mathcal{O}(d)$ immediately implies that $|a| = 1 + \mathcal{O}(d)$ and $\|b\| = \mathcal{O}(d)$. Then, following the discussion above, it suffices to bound $\|X_2T\|$ and the norms of the blocks in (4.5). The condition number of X_2T is bounded above by the product of these two estimates. Since $\|r(\tilde{\Lambda}_1)^{-1}\| = |r(\lambda_m)|^{-1}$, we obtain that

$$1/\sigma_m(X_2T) \leq \|(V_1^*X_2T)^{-1}\| \leq 1 + \frac{\|D^{-1}\|}{|r(\lambda_m)|}\left(1+\frac{\|b\|}{d}\right)+\mathcal{O}(d). \tag{4.7}$$

On the other hand, we can write V^*X_2T in block form, analogous to (4.4), as

$$V^*X_2T = \begin{bmatrix} a & b/(de^{i\theta}) \\ & r(\tilde{\Lambda})\tilde{D} \end{bmatrix} + \mathcal{O}(d),$$

where $\tilde{\Lambda} = \operatorname{diag}(\lambda_2, \ldots, \lambda_n)$ and $\tilde{D} = \tilde{V}^*\tilde{Q}_1$ (here, $\tilde{V}$ is V with the first column removed). Calculating the norm of the block diagonal component and the off-diagonal component separately and applying the triangle inequality yields $\|X_2T\| \leq M$. Combining this bound with the bound in (4.7) concludes the proof. ∎

Theorem 4.1 tells us that X_2 is only a simple column scaling away from a well-conditioned basis when the first column of Q_1 approximates v_1 with accuracy $\mathcal{O}(d)$. Since the sensitivity (and numerical computation) of the QR factorization is not affected by column scaling, the $\mathcal{O}(1)$ bound on $\kappa(X_2T)$ in (4.6) explains why the computed orthonormal basis for $\mathcal{S}_2$ is accurate to unit round-off. This line of analysis follows naturally from our observation about the eigenvector coordinates of $\hat{Q}_1$ in Figure 3.2 (right), but one question remains: Why is the first column of the computed orthonormal basis such a good approximation to v_1?

The answer is that $\hat{q}_1^{(1)}$ is essentially the first column of X_1 after normalization, up to the unit round-off u. In particular, $\hat{q}_1^{(1)}$ is unaffected by the u/d errors in $\hat{Q}_1$ caused by ill-conditioning in X_1 (see Proposition 3.1). These errors are concentrated in the latter columns of $\hat{Q}_1$ because of the nested structure of Householder reflections (or Givens rotations) used to make X_1 upper triangular. We have that $x_1/\|x_1\| = v_1 + \mathcal{O}(d)$ by (3.1), so we expect that $\hat{q}_1^{(1)} = v_1 + \mathcal{O}(d)$ also, as observed in Figure 3.2.

Finally, if the orthogonal factor is computed with modified Gram–Schmidt instead of Householder reflections or Givens rotations, the columns of $\hat{Q}_1$ lose orthogonality in proportion to the condition number of the ill-conditioned basis X_1. The consequence of this is that the block $v_1^*(\hat{Q}_1)_{(2:m)}$ from (4.1) may be as large as u/d instead of $\mathcal{O}(d)$, even though $\hat{q}_1^{(1)} = v_1 + \mathcal{O}(d)$. This may alter the order of magnitude of $\kappa(X_2T)$ when $d \ll \sqrt{u}$ (since then, $u/d \gg d$) as the balance in Theorem 4.1 is disrupted. In particular, twice may no longer be enough to correct ill-conditioning in $\hat{X}_2$. A similar effect is observed for nonnormal matrices in section 6 even when Householder reflections or Givens rotations are employed in the QR factorizations.

5. Convergence and stability. So far, our analysis of dangerous eigenvalues has focused on the conditioning of the iterates $X_1, X_2, \ldots$ in (2.1) and the corresponding accuracy in the computed orthonormal bases. Indeed, this perspective explains the u/d errors observed in the first iteration (see Figure 3.2) and provides essential insight into the restored accuracy observed in the second iteration (see Figure 4.1). But we have not yet explained how the round-off errors incurred while applying the ill-conditioned rational filter enter into the picture. Nor have we discussed how these round-off errors, together with the error in the computed orthonormal basis, accumulate during the iterations in (2.1).

To apply the rational filter $r(A)$ to an $n \times m$ matrix Q in practice, one solves linear systems with a shift at each pole and takes a weighted average of the solutions:

$$r(A)Q = \sum_{j=1}^{\ell} \omega_j X^{(j)}, \quad \text{where} \quad (z_j I - A)X^{(j)} = Q, \quad j = 1, \ldots, \ell. \tag{5.1}$$

If the linear systems are solved with a backward stable algorithm, then the computed solutions $\hat{X}^{(j)}$ satisfy, for each $j = 1, \ldots, \ell$,

$$(z_j I - A - \mathcal{E}_j)\hat{X}^{(j)} = Q, \qquad \|\mathcal{E}_j\| \le \gamma \|A\| u. \tag{5.2}$$

Here, $\mathcal{E}_j$ is the backward error and γ is a constant, with modest dependence on $z_1, \ldots, z_\ell$ and the dimension of Q, such that $\gamma u \ll 1$ for typical situations.[6]

Now, if we neglect errors made while forming the linear combination on the left-hand side of (5.1), then the forward error in $r(A)Q$ can be written as[7]

$$\hat{X} - r(A)Q = \sum_{j=1}^{\ell} \omega_j (z_j I - A)^{-1} \mathcal{E}_j \hat{X}^{(j)}, \qquad \text{where} \qquad \hat{X} = \sum_{j=1}^{\ell} \omega_j \hat{X}^{(j)}. \tag{5.3}$$

Due to the appearance of $\mathcal{E}_j$, the terms in the left-hand sum are all on the order of u except for the term corresponding to the pole near the dangerous eigenvalue, whose index we call $j = j_*$. In the dangerous term, $(z_{j_*} I - A)^{-1}$ amplifies the v_1 components in the columns of $\mathcal{E}_{j_*}$ by a factor of $1/d$. Similarly, the components of v_1 in the columns of Q are amplified to order $1/d$ in the corresponding columns of $\hat{X}^{(j_*)}$ (this is made precise by expanding (5.2) in a Neumann series). Therefore, the relative errors in the columns of $\hat{X}$ are on the order of u/d.

Thus, every time $r(A)$ is applied in (2.1), relative errors of order u/d are accrued in the columns of $\hat{X}_k$. On the one hand, our understanding of accuracy in the computed orthonormal basis $\hat{Q}_k$ (developed in sections 3 and 4) remains intact, because perturbations of relative order u/d to the columns of X_k have little effect on the leading order estimates for $\kappa(X_k)$. On the other hand, we may wonder: what effect do such perturbations have on $\operatorname{span}(X_k)$ and the geometric convergence implied in Theorem 2.1?

Recent analyses of subspace iteration accelerated with a rational filter suggest that $\operatorname{span}(\hat{X}_k)$ tends to $\mathcal{V}$ geometrically at roughly the expected rate until a threshold

[6] This characterization can be modified to accommodate inexact solution techniques, such as iterative methods, but γ may be much larger, depending on the stability properties of the particular numerical method [15, p. 339].

[7] For expositional clarity, we neglect round-off errors accrued when forming the linear combination in the right-hand side of (5.1) to focus on the effect of the ill-conditioned linear systems. For typical choices of the weights and nodes in $r(A)$, this amounts to discarding a term on the order of u relative to the largest column norm of the $\hat{X}^{(j)}$.

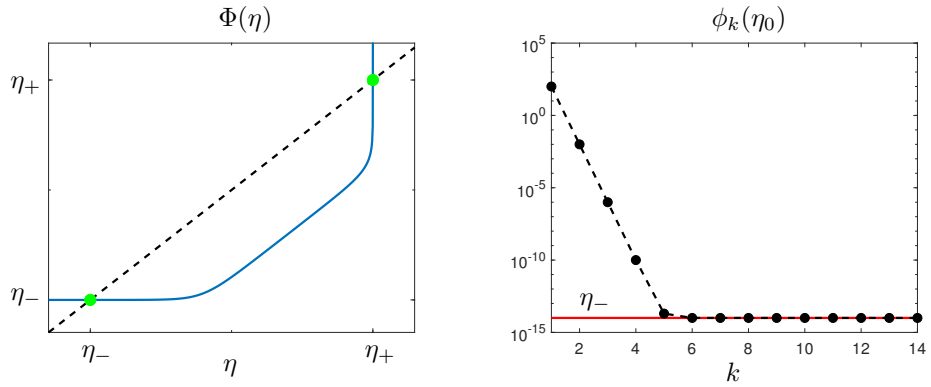

FIG. 5.1. *The dynamics of perturbed subspace iteration from* (5.14). *In the left panel, the solid line is the graph of* $\Phi(\eta)$, *and its fixed points (green circles) are marked at the intersections* $\Phi(\eta_\pm) = \eta_\pm$. *If* $\tan\theta_1(\hat{\mathcal{S}}_0, \mathcal{V})$ *falls between the two fixed points (green circles), then the* $\tan\theta_1(\hat{\mathcal{S}}_k, \mathcal{V})$ *must converge geometrically to a threshold near the lower fixed point (see Theorem* 5.4). *In the right panel, the iterated map* $\phi_k(\eta_0)$ *(circles) is compared with the upper bound in Theorem* 5.4 *(dashed line) for* $k = 0, \ldots, 14$. *For this experiment,* $\eta_0 = 100$, $\epsilon_1 = 10^{-5}$, $\epsilon_2 = 10^{-14}$, *and* $\rho = 10^{-4}$. *(See online version for color.)*

of accuracy is reached, at which point convergence plateaus [17]. This threshold is usually the same order of magnitude as the error accrued in the subspace at each iteration, i.e., in the columns of $\hat{X}_k$. Similar results have been derived for perturbations in the entries of the matrix $r(A)$ (this work does not consider filters explicitly) [13]. However, the evidence of the experiments in Figures 1.1 and 1.2 and in section 4 indicates that errors in $\mathrm{span}(\hat{X}_k)$ caused by dangerous eigenvalues do not prevent the Rayleigh–Ritz procedure from finding vectors in $\mathrm{span}(\hat{X}_k)$ that approximate the target eigenvectors to unit round-off accuracy. We now show that errors in $\hat{X}_k$ caused by the dangerous eigenvalue do not lead to early stagnation or instability in the computed iterates. In the worst case, they may slow the geometric convergence rate by a factor of roughly $(1 - u/d)^{-1}$. Moreover, the iteration is stable as long as the columns of the initial guess Q_0 are not too close to $\mathcal{V}^\perp$ (see Figure 5.1).

5.1. One-step refinement bounds. The amplifying power of the dangerous eigenvalue leads to large relative errors in the columns of $\hat{X}_k$. However, the errors possess an important quality: the amplification is entirely in the direction of v_1 so that the relative errors in the unwanted direction are still small. To understand how these structured perturbations influence $\hat{\mathcal{S}}_k = \mathrm{span}(\hat{X}_k)$, we gather the errors accrued during the kth iteration into a perturbation to the orthonormal basis for $\hat{\mathcal{S}}_{k-1}$ and construct a one-step refinement bound as in Theorem 2.3. To formulate this precisely, we replace (2.1) with the perturbed form

$$\hat{X}_k = r(A)(Q'_{k-1} + R_k), \qquad Q'_k = \mathrm{qf}(\hat{X}_k). \tag{5.4}$$

Note that we include any errors in the computed orthonormal factor in R_k, placing the emphasis on $\hat{\mathcal{S}}_k = \mathrm{span}(\hat{X}_k) = \mathrm{span}(Q'_k)$ rather than $\mathrm{span}(\hat{Q}_k)$. This causes no difficulty since, as we know from section 4, the error $\hat{Q}_k - Q'_k$ is on the order of u for $k \geq 2$. Since Q'_k is an orthonormal basis, $\hat{\mathcal{S}}_k$ and $\mathrm{span}(Q'_k)$ only differ by a term not much larger than u.

To begin, we establish the form (5.4) by way of the residuals of the linear systems in (5.3) and study the structure of R_k. To measure the columns of R_k relative to the

columns of $\hat{X}_k$, it is convenient to apply the diagonal scaling

$$C_k = \operatorname{diag}(\|(\hat{X}_k)_1\|^{-1}, \ldots, \|(\hat{X}_k)_m\|^{-1})/\sqrt{m}, \tag{5.5}$$

so that $\|\hat{X}_k C_k\| \leq 1$. We also need the majorization of the rational filter, denoted

$$\tilde{r}(\lambda) = \sum_{j=1}^{\ell} |\omega_j||(z_j - \lambda)^{-1}|. \tag{5.6}$$

As usual, $\tilde{r}(\Lambda_1)$ and $\tilde{r}(\Lambda_2)$ are the matrices when the function in (5.6) is applied to the diagonal matrices Λ_1 and Λ_2. Observe that $\|\tilde{r}(\Lambda_1)r(\Lambda_1)^{-1}\| = \mathcal{O}(1)$ as $d \to 0$, because the poles near the dangerous eigenvalue cancel.

LEMMA 5.1. *Let normal $A \in \mathbb{C}^{n\times n}$ and $r : \Lambda \to \mathbb{C}$ satisfy* (1.4) *and* (1.5)*, respectively. Given $Q \in \mathbb{C}^{n\times m}$, let $\hat{X} = \sum_{j=1}^{\ell} \omega_j \hat{X}^{(j)}$, with each $\hat{X}^{(j)}$ satisfying* (5.2)*. Then, there is an $R \in \mathbb{C}^{n\times m}$ such that $\hat{X} = r(A)(Q + R)$ and*

$$\|P_{\mathcal{V}} R\| \leq \gamma_1 \|A\| u/d \qquad \text{and} \qquad \|(I - P_{\mathcal{V}})r(A)RC\| \leq \gamma_2 \|A\| u. \tag{5.7}$$

Here, $\gamma_1 = \gamma\,\|r(\Lambda_1)^{-1}\tilde{r}(\Lambda_1)\|$, $\gamma_2 = \gamma\,\|\tilde{r}(\Lambda_2)\|$, and C is the diagonal scaling in (5.5) *(with index k suppressed).*

Proof. Because each $\hat{X}^{(j)}$ satisfies (5.2), and $\hat{X} = \sum_{j=1}^{\ell} \omega_j \hat{X}^{(j)}$, we collect like terms in (5.3) and compute

$$\hat{X} = r(A)Q + \sum_{j=1}^{\ell} \omega_j (z_j I - A)^{-1} R^{(j)}, \tag{5.8}$$

where $R^{(j)} = \mathcal{E}^{(j)}\hat{X}$. Note that $\|R^{(j)}C\| \leq \gamma\|A\|u$ for $j = 1, \ldots, \ell$, by (5.2).

We compute R directly by comparing (5.8) with $\hat{X} = r(A)(Q+R)$ and noting that we need $r(A)R = \sum_{j=1}^{\ell} \omega_j (z_j I - A)^{-1} R^{(j)}$. Inserting the eigenvalue decomposition $A = V\Lambda V^*$ into both sides and inverting $r(A) = V r(\Lambda) V^*$, we obtain

$$R = V r(\Lambda)^{-1} \left(\sum_{j=1}^{\ell} \omega_j (z_j I - \Lambda)^{-1} V^* R^{(j)} \right). \tag{5.9}$$

Calculating $P_{\mathcal{V}} R$ and $(I - P_{\mathcal{V}})r(A)RC$ directly from (5.9) and applying the backward error bounds in (5.2) to bound the residuals $\|R^{(j)}\|$ uniformly, we obtain the bounds in (5.7). ∎

Lemma 5.1 demonstrates that the perturbations R_k in (5.4) capture the essential structure of the errors in $\hat{X}_k$. First, R_k perturbs Q'_{k-1} with relative magnitude u/d and direction in the subspace $\mathcal{V}$. Second, $r(A)R_k$ perturbs the columns of X_k with relative magnitude u and direction in the subspace $\mathcal{V}^{\perp}$. We note that $\|V_2^* R_k C_k\|$ itself is not small when the filter is very good, i.e., close to unit round-off on the unwanted eigenvalues, as $\|r(\Lambda_2)^{-1}\tilde{r}(\Lambda_2)\|$ may be extremely large. However, the forward application of the filter cancels any large factors in $r(\Lambda_2)^{-1}$ exactly.

With Lemma 5.1 in hand, we can calculate a one-step refinement bound generalizing Theorem 2.3 to the perturbed iteration in (5.4). While the u/d relative errors in $\hat{X}_k$ are felt in the refinement factor in (5.10), they do not appear in the additive perturbation to $\tan\theta_1(\hat{\mathcal{S}}_k, \mathcal{V})$. This point is crucial because, as we show in subsection 5.2, the size of the additive term determines the threshold for stagnation in the worst-case accumulation of errors.

THEOREM 5.2. *Let normal $A \in \mathbb{C}^{n\times n}$ and $r : \Lambda \to \mathbb{C}$ satisfy* (1.4) *and* (1.5)*, respectively, and let $\hat{\mathcal{S}}_k = \operatorname{span}(\hat{X}_k)$, with $\hat{X}_k$ defined as in* (5.4) *and R_k satisfying* (5.7)*. If $\cos\theta_1(\hat{\mathcal{S}}_{k-1}, \mathcal{V}) > \gamma_1\|A\|u/d$ and $\cos\theta_1(\hat{\mathcal{S}}_k, \mathcal{V}) > 0$, then*

$$\tan\theta_1(\hat{\mathcal{S}}_k, \mathcal{V}) \le \Big|\frac{r(\lambda_{m+1})}{r(\lambda_m)}\Big| \frac{\tan\theta_1(\hat{\mathcal{S}}_{k-1}, \mathcal{V})}{1-\alpha_k} + \beta_k, \tag{5.10}$$

where $\alpha_k \le \gamma_1\|A\|u/(d\cos\theta_1(\hat{\mathcal{S}}_{k-1}, \mathcal{V}))$ and $\beta_k \le \gamma_2\|A\|\kappa(\hat{X}_kC_k)u/\cos\theta_1(\hat{\mathcal{S}}_k, \mathcal{V})$.

Proof. Calculating directly as in the proof of Theorem 2.3, we have that

$$T(\hat{X}_kC_k, V_1) = (I - P_{\mathcal{V}})r(A)(Q'_{k-1} + R_k)C_k(V_1^*\hat{X}_kC_k)^{-1}. \tag{5.11}$$

We proceed by bounding the two terms in (5.11) corresponding to Q'_{k-1} and R_k. By Lemma 5.1, $\|(I - P_{\mathcal{V}})r(A)R_kC_k\| \le \gamma_2\|A\|u$. If $\hat{X}_kC_k = Q'_kS_k$ is an economy-sized QR factorization, then the singular values of S_k and $\hat{X}_kC_k$ coincide, and

$$\|(V_1^*\hat{X}_kC_k)^{-1}\| = \|S_k^{-1}(V_1^*Q'_k)^{-1}\| \le \Big(\sigma_m(\hat{X}_kC_k)\cos\theta_1(\hat{\mathcal{S}}_k, \mathcal{V})\Big)^{-1}. \tag{5.12}$$

Since $\|\hat{X}_kC_k\| \le 1$, we conclude that $\|(I - P_{\mathcal{V}})r(A)R_kC_k(V_1^*\hat{X}_kC_k)^{-1}\| \le \beta_k$.

Now, rewrite $(V_1^*\hat{X}_k)^{-1} = (V_1^*(Q'_{k-1} + R_k))^{-1}r(\Lambda_1)^{-1}$ and expand

$$(V_1^*(Q'_{k-1} + R_k))^{-1} = \left(I + \sum_{j=1}^{\infty}(V_1^*Q'_{k-1})^{-j}(V_1^*R_k)^j\right)(V_1^*Q'_{k-1})^{-1}.$$

The Neumann series converges absolutely because $\|V_1^*R_k\| \le \gamma_1\|A\|u/d$ by Lemma 5.1 and $\|(V_1^*Q'_{k-1})^{-1}\| = (\cos\theta_1(\hat{\mathcal{S}}_{k-1}, \mathcal{V}))^{-1} < d/(\gamma_1\|A\|u)$ by hypothesis; consequently, $\|(V_1^*Q'_{k-1})^{-1}V_1^*R_k\| < 1$. Since $T(Q'_{k-1}, V_1) = (I - P_{\mathcal{V}})Q'_{k-1}(V_1^*Q'_{k-1})^{-1}$, we have

$$\begin{aligned}(I - P_{\mathcal{V}})r(A)Q'_{k-1}(V_1^*\hat{X}_k)^{-1} &= r(A)(I - P_{\mathcal{V}})Q'_{k-1}(V_1^*(Q'_{k-1} + R_k))^{-1}r(\Lambda_1)^{-1}\\ &= r(A)T(Q'_{k-1}, V_1)\left(I + \sum_{j=1}^{\infty}(V_1^*Q'_{k-1})^{-j+1}(V_1^*R_k)^j(V_1^*Q'_{k-1})^{-1}\right)r(\Lambda_1)^{-1}.\end{aligned}$$

Because the range of $T(Q'_{k-1}, V_1)$ is $\mathcal{V}^\perp$, the first factor on the right-hand side is bounded by $\|r(A)T(Q'_{k-1}, V_1)\| \le \|r(\Lambda_2)\|\tan\theta_1(\hat{\mathcal{S}}_{k-1}, \mathcal{V})$. The factor in parentheses is bounded above by $\sum_{j=0}^{\infty}\alpha_k^j = (1-\alpha_k)^{-1}$, where $\alpha_k = \|(V_1^*Q'_{k-1})^{-1}\|\|V_1^*R_k\|$. Therefore, we have the upper bound

$$\|(I - P_{\mathcal{V}})r(A)Q'_{k-1}(V_1^*\hat{X}_k)^{-1}\| \le \|r(\Lambda_2)\|\|r(\Lambda_1)^{-1}\|\frac{\tan\theta_1(\hat{\mathcal{S}}_{k-1}, \mathcal{V})}{1-\alpha_k}.$$

Noting that $\|r(\Lambda_2)\| = |r(\lambda_{m+1})|$, $\|r(\Lambda_1)^{-1}\| = |r(\lambda_m)|^{-1}$ and collecting the bounds for the two terms in (5.11) establishes (5.10). ∎

The significance of Theorem 5.2 is that the errors in $\hat{X}_k$ that lie in the target subspace $\mathcal{V}$ impact only the refinement rate and do not contribute to the additive term β_k in (5.10). This worst-case scenario occurs over one iteration only when the perturbations are aligned to maximally cancel the components of $\mathcal{V}$ present in the basis Q'_{k-1}. In fact, such errors are just as likely to align perfectly with the $\mathcal{V}$ components of

Q'_{k-1} and improve the refinement rate by $(1+\alpha_k)^{-1}$, so the impact on the geometric convergence rate implied by (5.10) is probably not observed in practice.

On the other hand, the errors in $\hat{X}_k$ that lie in $\mathcal{V}^\perp$ degrade the expected refinement through the additive term β_k and, due to orthogonality, have a tangible effect on the convergence of subspace iteration in floating-point arithmetic. Note that the magnitude of β_k is proportional to the condition number of the basis $\hat{X}_k$ after column scaling. From sections 3 and 4, we know that $\beta_1 \approx u/d$ and $\beta_k \approx u$ for $k \geq 2$, provided that $\hat{\mathcal{S}}_k$ and $\mathcal{V}$ do not become too close to orthogonal during the iteration.

5.2. Stability and stagnation. According to Theorem 5.2, the search subspace is refined by a factor comparable to Theorem 2.1, up to the size of the errors β_k introduced in $\mathcal{V}^\perp$, at each iteration. As we accumulate iterations, the errors in $\mathcal{V}^\perp$ are filtered out by $r(A)$ and, in the apt words of the authors of [17], "the dominant error term is the one most recently introduced." As long as $\cos\theta_1(\hat{\mathcal{S}}_k, \mathcal{V})$ is bounded sufficiently far from zero for $k \geq 0$, the sequences α_k and β_k remain stable at the orders of u/d and u (respectively) after the first iteration. In this case, we expect that $\hat{\mathcal{S}}_k$ converges geometrically toward $\mathcal{V}$ until a threshold of about u is reached, after which convergence stagnates. This is what we observe in Figures 1.1 and 1.2.

If $\cos\theta_1(\hat{\mathcal{S}}_k, \mathcal{V})$ does become very small at some step in the iteration, then the one-step refinement bound may not imply any refinement in the search subspace at all: the iteration in (5.4) is potentially unstable. With a slight change of perspective, we now characterize the behavior of the iterates in (5.4) as $k \to \infty$, addressing both the stability and the threshold for stagnation in subspace refinement.

Let us introduce the constants $\rho = |r(\lambda_{m+1})|/|r(\lambda_m)|$, $\epsilon_1 = \gamma_1\|A\|u/d$, and $\epsilon_2 = \gamma_2\|A\|\hat{M}u$, where $\hat{M}$ is an $\mathcal{O}(1)$ uniform bound on $\kappa(X_kC_k)$ for $k \geq 2$ (i.e., from Theorem 4.1). Consider the function

$$\Phi(\eta) = \frac{1}{1-\epsilon_2}\left(\frac{\rho\,\eta}{1-\epsilon_1(1+\eta)} + \epsilon_2\right). \tag{5.13}$$

Because $1/\cos\theta \leq 1 + \tan\theta$ when $0 \leq \theta \leq \pi/2$, we can rewrite Theorem 5.2 in the form $\tan\theta_1(\hat{\mathcal{S}}_k, \mathcal{V}) \leq \Phi(\tan\theta_1(\hat{\mathcal{S}}_{k-1}, \mathcal{V}))$ when $k \geq 2$. We can understand the "worst-case" behavior of subspace iteration by studying the trajectory of $\tan\theta_1(\hat{\mathcal{S}}_0, \mathcal{V})$, for some initial subspace $\hat{\mathcal{S}}_0$, obtained by iterating the map Φ.

Given $\eta_0 > 0$, let $\phi_k(\eta_0)$ denote the k-fold iteration of the map Φ on the point η_0, so that (letting $f \circ g$ denote the composition of two functions) we have

$$\phi_k(\eta_0) = \underbrace{\Phi \circ \cdots \circ \Phi}_{k}(\eta_0). \tag{5.14}$$

We call η_* a fixed point of Φ if $\Phi(\eta_*) = \eta_*$ and say that η_* is monotone attracting for $\Omega \subset [0, \infty)$ if $\phi_k(\eta) \to \eta_*$ monotonically as $k \to \infty$ for all $\eta \in \Omega$. After applying Φ to $\tan\theta_1(\hat{\mathcal{S}}_0, \mathcal{V})$ k times, we see that (since $\Phi(\eta)$ is nondecreasing on $0 \leq \eta < -1 + 1/\epsilon_1$)

$$\tan\theta_1(\hat{\mathcal{S}}_k, \mathcal{V}) \leq \phi_k\left(\tan\theta_1(\hat{\mathcal{S}}_0, \mathcal{V})\right) \tag{5.15}$$

as long as $\phi_j(\tan\theta_1(\hat{\mathcal{S}}_0, \mathcal{V})) < -1 + 1/\epsilon_1$ for each $j \geq 1$. Consequently, the fixed points of Φ and their attracting sets provide insight into the behavior of $\tan\theta_1(\hat{\mathcal{S}}_k, \mathcal{V})$ in the limit $k \to \infty$, that is, the convergence and stability of the iteration in (5.4).

LEMMA 5.3. *Define the map* $\Phi : [0, -1 + 1/\epsilon_1) \to [0, \infty)$ *as in* (5.13)*, with constants* $0 < \rho < 1$ *and* $0 < \epsilon_1,\ \epsilon_2 < 1$*. Let*

$$\delta = \frac{1}{2\epsilon_1}\left[1 - \frac{\rho}{1-\epsilon_2} - \epsilon_1\left(1 - \frac{\epsilon_2}{1-\epsilon_2}\right)\right] \quad \textit{and} \quad \sigma = \frac{\epsilon_2(1-\epsilon_1)}{\epsilon_1(1-\epsilon_2)}.$$

If $\delta^2 > \sigma$*, then* Φ *has precisely two fixed points, given by* $\eta_\pm = \delta \pm \sqrt{\delta^2 - \sigma}$*. Moreover, the fixed point* η_- *is monotone attracting on* $[0, \eta_+)$*.*

Proof. Starting from the fixed point equation $\Phi(\eta_*) = \eta_*$, we multiply through by $(1 - \epsilon_1(1 + \eta_*))$ and collect powers of η_* to obtain the quadratic equation

$$\epsilon_1\eta_*^2 - \left[1 - \frac{\rho}{1-\epsilon_2} - \epsilon_1\left(1 - \frac{\epsilon_2}{1-\epsilon_2}\right)\right]\eta_* + \frac{\epsilon_2(1-\epsilon_1)}{1-\epsilon_2} = 0. \tag{5.16}$$

Applying the quadratic formula for the roots and rewriting in terms of δ and σ concludes the fixed-point calculation. Now, the quadratic on the left-hand side of (5.16) is negative between the roots, which implies that $\Phi(\eta) > \eta$ for $0 < \eta < \eta_-$ and $\Phi(\eta) < \eta$ for $\eta_- < \eta < \eta_+$. This change at each fixed point implies that η_- attracts nearby points and that η_+ repels nearby points. Because Φ is nondecreasing and has no other fixed points, we conclude that η_- is monotone attracting on $[0, \eta_+)$. ∎

Lemma 5.3 shows that if $\tan\theta_1(\hat{\mathcal{S}}_0, \mathcal{V}) < \eta_+$, then $\tan\theta_1(\hat{\mathcal{S}}_k, \mathcal{V})$ must eventually be on the order of η_- or better for all sufficiently large k. Recalling that the constants ϵ_1 and ϵ_2 are on the order of u/d and u, respectively, and that ρ is the filtered spectral ratio, we estimate the size of the fixed points to be

$$\eta_- \approx \frac{\epsilon_2}{1-\rho} \quad \text{and} \quad \eta_+ \approx -1 + \frac{1-\rho}{\epsilon_1}. \tag{5.17}$$

Crucially, the lower fixed point η_- is on the order of u, not u/d. Having established stability properties of the perturbed iteration in (5.4), we can now estimate the rate of convergence to the fixed point η_-.

THEOREM 5.4. *Define the map* $\Phi : [0, -1 + 1/\epsilon_1) \to [0, \infty)$ *as in* (5.13)*, with constants* $0 < \rho < 1$ *and* $0 < \epsilon_1,\ \epsilon_2 < 1$*. Let* ϕ_k *denote the* k*-fold iteration of* Φ *as in* (5.14)*. If* Φ *satisfies the hypotheses of Lemma* 5.3*, then given* $0 \le \eta_0 < \eta_+$*, it holds that*

$$\phi_k(\eta_0) \le \tilde{\rho}^k\eta_0 + \tilde{\epsilon}_2(1-\tilde{\rho})^{-1}, \qquad k \ge 1. \tag{5.18}$$

Here, $\tilde{\rho} = \rho(1-\epsilon_2)^{-1}(1 - \epsilon_1(1+\eta_0))^{-1}$ *and* $\tilde{\epsilon}_2 = \epsilon_2(1-\epsilon_2)^{-1}$*.*

Proof. Denote $\eta_k = \phi_k(\eta_0)$ for each $k \ge 1$. From the definitions of Φ and ϕ_k in (5.13) and (5.14), respectively, we compute that

$$\eta_k = \Phi(\eta_{k-1}) = \tilde{\rho}\eta_{k-1} + \tilde{\epsilon}_2, \qquad k \ge 1. \tag{5.19}$$

By hypothesis, Lemma 5.3 applies, so $\eta_k \to \eta_-$ monotonically as $k \to \infty$ and, consequently, $\tilde{\rho} < 1$. Therefore, we iterate (5.19) $k-1$ times to obtain

$$\eta_k = \tilde{\rho}^k\eta_0 + \tilde{\epsilon}_2\sum_{j=0}^{k-1}\tilde{\rho}^j \le \tilde{\rho}^k\eta_0 + \tilde{\epsilon}_2/(1-\tilde{\rho}).$$

Plugging the original parameters back into $\tilde{\rho}$ and $\tilde{\epsilon}_2$ establishes (5.18). ∎

Thus, Theorem 5.4 and (5.15) demonstrate that the reduction of $\tan\theta_1(\hat{\mathcal{S}}_k, \mathcal{V})$ down to the order of η_- is approximately geometric with rate close to ρ. So (accounting for the fact that the additive perturbation term is actually on the order of u/d in the first iteration) it takes approximately $1 + \log(\eta_-)/\log(\rho)$ steps for $\hat{\mathcal{S}}_k$ to converge to within order u of $\mathcal{V}$, as measured by the tangent of the principal angle between the two subspaces.

6. Nonnormal matrices. We now consider the case of an $n \times n$ diagonalizable matrix A whose eigenvectors are not orthogonal. Although a straightforward extension of Proposition 3.1 shows that the condition number of X_1 still scales, generically, like $1/d$ (see Proposition 6.1 below), the effect of a dangerous eigenvalue on subsequent iterates, $X_2, X_3, \ldots$, computed via (2.1) is distinct in the nonnormal case due to interactions among nonorthogonal modes. In fact, the condition numbers of the computed iterates do not improve during subsequent iterations unless approximate eigenvectors (i.e., from Ritz vectors) are incorporated into the subspace iteration (see Algorithm 6.1). Even with this modification, the condition numbers may remain large after one iteration when d is very small (loosely, when $d \ll \sqrt{u}$), unlike the normal case. Here, we demonstrate that $\kappa(X_k)$ is typically reduced in step with the error in the Ritz vectors and that $\kappa(X_k) \approx (u/d)^k$ in the best case (i.e., when $|r(\lambda_{m+1})|/|r(\lambda_m)| \approx u$ and the Ritz vectors are well-conditioned at each iteration).

When A does not have an orthogonal basis of eigenvectors (but is still diagonalizable), the orthogonal spectral projectors $v_i v_i^*$ that diagonalize the filter in (3.1) are replaced by oblique spectral projectors, so that

$$r(A)x = \sum_{i=1}^{n} r(\lambda_i)\frac{w_i^* x}{w_i^* v_i} v_i = \frac{w_1^* x}{(de^{i\theta})(w_1^* v_1)} v_1 + \mathcal{O}(1) \quad \text{as} \quad d \to 0. \tag{6.1}$$

Here, $w_1, \ldots, w_n$ are the left eigenvectors of A, satisfying $w_i^* A = \lambda_i w_i^*$ with $\|w_i\| = 1$ for $i = 1, \ldots, n$. Likewise, the spectral decomposition in (1.4) is replaced by

$$A = V_1 \Lambda_1 W_1^* + V_1 \Lambda_2 W_2^*, \tag{6.2}$$

where the ith column of $W = [W_1\ W_2]$ is $(w_i^* v_i)^{-1} w_i$. With this normalization, V and W form a biorthogonal system, meaning that $W^* V = I$, with I being the $n \times n$ identity matrix. In the biorthogonal system, the dangerous eigenvalue amplifies the w_1 component in the input x along the v_1 direction in the output $r(A)x$. Due to biorthogonality, v_1 and w_1 are parallel only when v_1 is orthogonal to $v_2, \ldots, v_n$.

6.1. First iteration. To develop a sense of how nonnormality impacts the conditioning of the iterates, it is worthwhile to revisit the analysis of $\kappa(X_1)$ in Proposition 3.1 when A is only diagonalizable. While the condition number of X_1 is still $\mathcal{O}(1/d)$ as $d \to 0$, the constants in the bound now depend on the structure of the left and right eigenvectors. This is because the stretching and shrinking actions of A no longer belong solely to its eigenvalues but can be enhanced or attenuated by interactions among nonorthogonal eigenvectors. We denote the smallest singular values of V_1 and W_1 by $\sigma_m(V_1)$ and $\sigma_m(W_1)$, respectively.

PROPOSITION 6.1. *Let diagonalizable $A \in \mathbb{C}^{n\times n}$ and $r : \Lambda \to \mathbb{C}$ satisfy* (6.2) *and* (1.5)*, respectively, and given orthonormal $Q_0 \in \mathbb{C}^{n\times m}$, let $X_1 = r(A)Q_0$. If $U_1 = \mathrm{qf}(W_1)$ and $U_1^* Q_0$ has full rank, then the condition number of X satisfies*

$$\frac{\|w_1^* Q_0\|/|w_1^* v_1|}{d\kappa(V)|r(\lambda_2)|} \lesssim \kappa(X_1) \leq \left|\frac{r(\lambda_1)}{r(\lambda_m)}\right| \frac{\kappa(V)\|(U_1^* Q_0)^{-1}\|}{\sigma_m(V_1)\sigma_m(W_1)} \quad \text{as} \quad d \to 0. \tag{6.3}$$

Proof. The steps of the proof are essentially identical to those in Proposition 3.1 if (6.1) and (6.2) are used in place of (1.4) and (3.1), so we emphasize the adaptations made for nonorthogonal eigenvectors. For the largest singular value of X_1, we bound $\sigma_1(X_1) = \|r(A)Q_0\| \leq \kappa(V)|r(\lambda_1)|$, since $|r(\lambda_1)| \leq \|r(A)\| \leq \kappa(V)\|r(\Lambda)\|$ in the nonnormal case. If we use (6.2) to decompose X_1 as in (3.5), the singular values of $r(A)W_1^*Q_0$ do not tell us directly about the singular values of X_1 because V is not unitary. However, if $\Omega_1 R_1 = V_1$ and $\Omega_2 R_2 = V_2$ are economy-sized QR factorizations, we can decompose

$$r(A)Q_0 = \begin{bmatrix}\Omega_1 & \Omega_2\end{bmatrix} \begin{bmatrix}R_1 r(\Lambda_1) W_1^* Q_0 \\ R_2 r(\Lambda_2) W_2^* Q_0\end{bmatrix}.$$

Since Ω_1 and Ω_2 have orthonormal columns, we apply the argument in the proof of Proposition 3.1 to obtain the bound $1/\sigma_m(X_1) \leq \|(R_1 r(\Lambda_1) W_1^* Q_0)^{-1}\|$. Now, R_1 has the same singular values as V_1 and $\|R_1^{-1}\| = 1/\sigma_m(R_1)$, so we have that

$$\kappa(X_1) \leq \frac{|r(\lambda_1)|}{|r(\lambda_m)|} \frac{\kappa(V)\|(W_1^* Q_0)^{-1}\|}{\sigma_m(V_1)}. \tag{6.4}$$

The upper bound in (6.3) follows by substituting the QR decomposition $U_1 S_1 = W_1$ into (6.4) and noting that $\|S_1^{-1}\| = 1/\sigma_m(W_1)$.

A lower bound on $\sigma_1(X_1)$ follows directly from (6.1), which is analogous to (3.6). For the lower bound on $1/\sigma_m(X_1)$, we can use (6.1) to write X_1 as a rank one perturbation of the matrix

$$\tilde{N}_2 = V \mathrm{diag}(0, \lambda_2, \ldots, \lambda_n) W^* Q_0.$$

We have that $\sigma_1(N_2) \leq \|V\|\|W^*\||r(\lambda_2)| = \kappa(V)|r(\lambda_2)|$, where the equality is due to biorthogonality, which implies that $W^* = V^{-1}$. By interlacing, we find that $1/\sigma_m(X_1) \geq 1/(\kappa(V)|r(\lambda_2)|)$, establishing the asymptotic lower bound in (6.3). ∎

When A is normal, Proposition 6.1 reduces to Proposition 3.1. In the nonnormal case, ill-conditioning in the eigenvectors, reflected in $\kappa(V)$, widens the interval between the upper and lower bounds. Similarly, ill-conditioning in the target eigenvectors, captured by the smallest singular values of V_1 and W_1 (since the columns of both matrices have unit norm) may further widen the gap. On the other hand, the dangerous eigenvalue itself is ill-conditioned when $|w_1^* v_1|$ is small.[8] The left-hand side of (6.3) illustrates how this may enhance the amplifying effects of the dangerous eigenvalue, increasing the asymptotic lower bound to $d|w_1^* v_1|^{-1}$. Broadly speaking, the widening gap between upper and lower bounds indicates that our picture is blurred in the nonnormal case because the structure of the eigenvectors plays a key role. The extent of the damage may depend on where the ill-conditioning in V is concentrated.

6.2. Iterating with orthonormal bases. Now that we understand the interaction between nonnormality and dangerous eigenvalues in the initial iteration, we are ready to examine subsequent iterations. As in section 4, we focus on the coordinates of Q_1 in the eigenvector basis, partitioned into blocks as

$$W_1^* Q_1 = \begin{bmatrix} w_1^* q_1^{(1)} & w_1^* \tilde{Q}_1 \\ \tilde{W}_1^* q_1^{(1)} & \tilde{W}_1^* \tilde{Q}_1 \end{bmatrix}. \tag{6.5}$$

[8] With $\|v_i\| = \|w_i\| = 1$, the quantity $|w_i^* v_i|^{-1}$ is Wilkinson's condition number for λ_i, measuring the first-order sensitivity of the eigenvalue to infinitesimal perturbations in A [22, pp. 88–89].

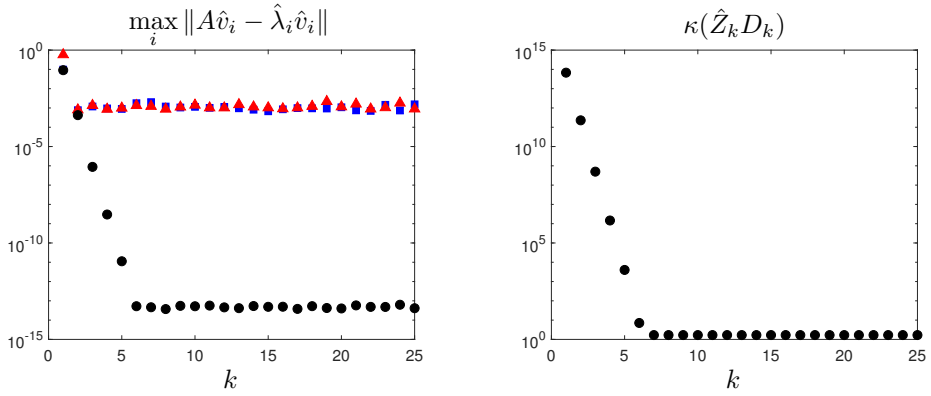

FIG. 6.1. *Dangerous eigenvalues of a nonnormal matrix. The eigenvalues and rational filter are identical to the setup displayed in Figure* 3.1*; however, this matrix has nonorthogonal eigenvectors, and the dangerous eigenvalue has been moved to distance* $d = 10^{-13}$ *from the pole at* $z = 10$*. On the left is shown the maximum residual of* 10 *target eigenpairs after each iteration of* (2.1) *(blue squares), a variant of subspace iteration based on Schur vectors* [12, Chap. 5.2] *(red triangles), and a variant based on approximate eigenvectors, described in Algorithm* 6.1 *(black circles). On the right, the condition number of the iterates* $\hat{Z}_k D_k$ *(*D_k *scales the columns of* $\hat{Z}_k$ *to have unit norm) decreases in step with residuals from Algorithm* 6.1 *at a rate of about* u/d *per iteration. (See online version for color.)*

The critical observation about (6.5) is that, in contrast to the normal case, the upper-right block is not small (the lower-left block remains small). Although the columns of $\tilde{Q}_1$ are still nearly orthogonal to v_1, the eigenvectors v_1 and w_1 are only parallel in the special case where v_1 is orthogonal to $v_2, \ldots, v_n$. Consequently, $w_1^* \tilde{Q}_1$ is typically $\mathcal{O}(1)$, and, when we compute $X_2 = r(A)Q_1$, the components in each column of Q_1 in the w_1 direction will be amplified according to (6.1). Each column of X_2 will be dominated by v_1 at magnitude $\mathcal{O}(1/d)$, and X_2 is just as ill-conditioned as X_1 in the first iteration. This line of thinking seems to indicate that, when $r(A)$ is repeatedly applied to an orthonormal basis, subspace iteration for nonnormal matrices must stagnate at an accuracy of $\approx u/d$ due to ill-conditioning in the iterates $X_1, X_2, \ldots$.

To illustrate, we return to the experimental setup illustrated in (3.1). We select the same rational filter and a matrix with the same eigenvalues, but now the eigenvector matrix is not orthogonal. The condition number of the eigenvector matrix is $\approx 10^2$, but the target eigenvectors themselves are not far from orthogonal. Figure 6.1 shows the maximum residual of the computed target eigenpairs after each of the first 10 iterations of (2.1). We also compare with a modified subspace iteration based on Schur vectors that is commonly used to compute eigenvalues of nonnormal matrices [12, Chap. 5.2]. Both iterations apply the rational filter directly to an orthonormal basis for the search space, and the residuals stagnate near u/d in both cases.

6.3. Iterating with approximate eigenvectors. What can we do to improve the conditioning of the iterates and the accuracy in the target eigenpairs? Consider another common variant of subspace iteration shown in Algorithm 6.1, which forms the iterates $Z_1, Z_2, \ldots$ by applying $r(A)$ to approximate eigenvectors constructed from the Ritz vectors at each iteration. Let us partition $W_1^* Y_1$ in the usual way,

$$W_1^* Y_1 = \begin{bmatrix} w_1^* y_1^{(1)} & w_1^* \tilde{Y}_1 \\ \tilde{W}_1^* y_1^{(1)} & \tilde{W}_1^* \tilde{Y}_1 \end{bmatrix} = \begin{bmatrix} e & f \\ g & H \end{bmatrix}, \tag{6.6}$$

Algorithm 6.1. Filtered subspace iteration with Rayleigh–Ritz projection.

Input: Given $A \in \mathbb{C}^{n\times n}$, $r : \Lambda \to \mathbb{C}$, and $Y_0 \in \mathbb{C}^{n\times m}$.

1: **for** $k = 1, 2, \ldots$ **do**
2: Apply the filter $Z_k = r(A)Y_{k-1}$.
3: Compute orthonormal basis $Q_k = \mathrm{qf}(Z_k)$.
4: Form $A_k = Q_k^* A Q_k$ and diagonalize $A_k = U_k \Theta_k U_k^{-1}$.
5: Set $Y_k = Q_k U_k$.
6: **end for**

Output: Approximate eigenvalue matrix Θ_k and eigenvector matrix Y_k.

where $\tilde{W}_1$ and $\tilde{Y}_1$ denote the last $m-1$ columns of W_1 and Y_1, respectively. Now, because the left and right eigenvectors are biorthogonal, w_1^* annihilates the remaining target eigenvectors $v_2, \ldots, v_m$, so the upper-right block f in (6.6) is small when the columns of Y_1 are a good approximation to the target eigenvectors. In turn, small $\|f\|$ mitigates the amplification of v_1 in the last $m-1$ columns of Z_2.

Unfortunately, the behavior of approximate eigenvectors computed with (6.1) may vary widely for general nonnormal matrices. In exact arithmetic, their accuracy will depend on the rational filter through the eigenvalues of $r(A)$ and on interactions among nonorthogonal eigenvectors. In floating-point arithmetic, their accuracy is further limited by the accuracy in the computed orthonormal basis and Ritz vectors. Despite these difficulties, we can glean some practical insight into a distinct feature of the nonnormal setting by examining a "best-case" situation.

Let us suppose that the nonnormal effects are relatively mild, that $r(\cdot)$ filters out the unwanted eigenvalues to unit round-off or better (as in Figure 3.1), and that the Ritz vectors are computed accurately at each iteration. In this regime, the accuracy of the approximate eigenvectors Y_1 is limited mainly by the accuracy in the computed orthonormal basis, $\hat{Q}_1$, and we can focus on the influence of the dangerous eigenvalue in the second iteration (and beyond). From our analysis of the first iteration in subsection 6.1, we expect that $\|\hat{Q}_1 - Q_1\| \approx u/d$ and therefore (by our assumptions on the filter and the Ritz vectors) that $\|\hat{Y}_1 - V_1\| \approx u/d$.

Interestingly, the order of magnitude of block f in (6.6) is distinctly different from the analogous block b in the normal case. Instead of the perfect balancing between b and $r(\lambda_1)$ when the filter is applied (leading to perfectly well-conditioned columns of X_2), we have the order-of-magnitude estimate $\|f\||r(\lambda_1)| \approx u/d^2$. In other words, v_1 may still dominate each column of Z_2 when $d \ll \sqrt{u}$, but the gap in magnitude between the v_1 component and the remaining target components in the last $m-1$ columns is reduced by a factor of u/d at the second iteration. Figure 6.1 illustrates this phenomenon in action, with the same matrix and rational filter used for the experiments in subsection 6.2. The residuals in the target eigenpairs decrease geometrically with rate u/d (left panel), mirroring the reduction in the condition number of the iterates Z_k (after scaling columns to have unit norm (right panel)).

Thus, for a mildly nonnormal matrix with a dangerous eigenvalue at distance $d \ll \sqrt{u}$ from a pole of $r(\cdot)$, two iterations are not usually enough to remove the adverse influence of the dangerous eigenvalue. Instead, the target residuals and the errors in the computed orthonormal basis are often refined in step down to the unit round-off. As in the normal case, round-off errors caused by the dangerous eigenvalue may even go unnoticed when the rational filter is mediocre, so that the noise in the unwanted directions is dominated by poor filtering.

7. Restarting Arnoldi. Now that we understand the right-hand side of Figure 1.1, let us examine the stagnation of Arnoldi with shift-and-invert enhancement, illustrated in the left-hand panel of the same figure. Unlike subspace iteration, which applies $r(A)$ iteratively to a subspace of fixed dimension, Arnoldi refines the subspace by expanding it. Given an initial unit vector $q_1 \in \mathbb{C}^n$, shift-and-invert Arnoldi computes the iterates

$$y_k = s(A)q_{k-1}, \qquad q_k = \text{mgsr}(y_k\,;\, q_1, \ldots, q_{k-1}), \tag{7.1}$$

with the expression mgrs$(\cdot)$ indicating that y_k is orthogonalized against $q_1, \ldots, q_{k-1}$ using modified Gram–Schmidt with full reorthogonalization [15, pp. 307–308].

After k steps of (7.1), we have an $n \times k$ orthonormal basis $Q_k = [q_1 \cdots q_k]$, and we can approximate eigenpairs of A in one of the following two ways:

- Directly from the eigenpairs of the upper Hessenberg matrix H_k generated from the weights calculated during modified Gram–Schmidt [20, p. 253].
- A Rayleigh–Ritz step by computing eigenpairs of $A_k = Q_k^* A Q_k$.

Usually, the upper Hessenberg matrix is the method of choice because it does not require any additional matrix-vector products. However, when a dangerous eigenvalue is present, the upper Hessenberg matrix in the Arnoldi decomposition of $s(A)$ typically has norm $\|H_k\| = \mathcal{O}(d^{-1})$: this makes the accurate calculation of the remaining target eigenvalues challenging for standard dense solvers. To focus on the accuracy in the computed basis Q_k, we work with A_k, but we revisit H_k at the end of this section.

In keeping with the analysis in sections 3 and 4, we can understand the accuracy in the computed orthonormal basis $\hat{Q}_k$ through the conditioning of the matrix

$$Y_k = \begin{bmatrix} q_1 & \cdots & q_{k-1} & y_k \end{bmatrix}, \qquad k = 2, 3, 4, \ldots. \tag{7.2}$$

The matrix Q_k from the Arnoldi iterations is precisely the QR factorization of Y_k obtained by orthogonalizing y_k against the previous $(k-1)$ columns, which are already an orthonormal set. If y_k is not too closely aligned with $\text{span}(q_1, \ldots, q_{k-1})$, then the matrix Y_k is well-conditioned, at least after a simple column scaling. Consequently, $Q_k = \text{qf}(Y_k)$ is not too sensitive to perturbations caused by round-off in Y_k, as discussed in subsection 3.1. However, if y_k is closely aligned with any of the previous columns, the smallest singular value of Y_k will be close to zero, and Q_k will be very sensitive to round-off in Y_k.

This perspective provides an explanation for the stagnation observed in Figure 1.1. When q_1 is chosen randomly, $v_1^* q_1$ is generically $\mathcal{O}(1)$ (as $d \to 0$). After applying the shift-and-invert filter, we calculate (as usual) that

$$y_2 = s(A)q_1 = \frac{v_1^* q_1}{d e^{i\theta}} v_1 + \mathcal{O}(1).$$

After we orthogonalize y_2 against q_1 to compute q_2, for some constant h_2 we have

$$q_2 = h_2 \frac{v_1^* q_1}{d e^{i\theta}} (v_1 - (v_1^* q_1) q_1) + \mathcal{O}(1). \tag{7.3}$$

In other words, q_2 may not be dominated by v_1, but $\text{span}(q_1, q_2)$ contains approximations to v_1 that are accurate to $\mathcal{O}(d)$.

Now, note that q_2 is not near orthogonal to v_1 unless q_1 happens to be very closely aligned with v_1. This means that the subsequent iterate y_3 is also aligned with v_1, and therefore with a vector in $\text{span}(q_1, q_2)$, to about order d. Consequently, the matrix Y_3

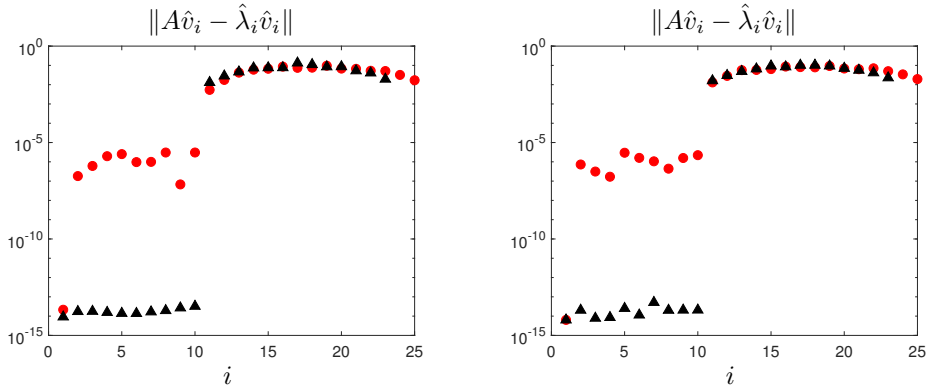

FIG. 7.1. *After restarting Arnoldi with Ritz vectors that are nearly orthogonal to the dangerous direction, Arnoldi produces approximations to the target eigenpairs with accuracy near the unit round-off. Both plots compare eigenpair residuals after* 25 *steps of shift-and-invert Arnoldi with no restart (red circles) to eigenpair residuals obtained after* 25 *total steps of shift-and-invert Arnoldi with the Ritz restart. The eigenpairs were extracted from* $Q_{25}^* A Q_{25}$ *in the left panel and from the Hessenberg matrix* H_{25} *in the right panel. (See online version for color.)*

is ill-conditioned and we expect that Q_3 and, in particular, q_3, can only be accurate to about order u/d when computed in floating-point precision. Moreover, q_3 is not dominated by v_1, and this process repeats, so that each iterate y_k is closely aligned with v_1 in $\text{span}(q_1, q_2)$, leading to errors in q_k on the order of u/d.

In our discussion above, note that y_3 was only aligned with v_1, and thus close to $\text{span}(q_1, q_2)$, because q_2 was not nearly orthogonal to v_1. Unlike in subspace iteration, the dangerous direction is never rendered harmless by orthogonalizing directly against it! The geometric picture of the iterates $y_2, y_3, y_4, \ldots$ being attracted to v_1 as a result of $q_2, q_3, q_4, \ldots$ not being sufficiently orthogonal to v_1 suggests an interesting fix. If we restart the Arnoldi iteration with the Ritz approximation associated to v_1 after the second iteration, the picture changes drastically. Again, y_2 is aligned with v_1, but now it is orthogonalized against $q_1 = v_1 + \mathcal{O}(d)$. The corresponding q_2 may not be particularly accurate, but this doesn't matter much: the point is that all subsequent iterates are orthogonalized against the dangerous direction (via q_1) up to order $\mathcal{O}(d)$. Analogous to the situation encountered in subspace iteration, the iterates $y_3, y_4, y_5, \ldots$, are no longer dominated by v_1. and, consequently, $q_3, q_4, q_5, \ldots$ can be computed accurately. In a sense, we are tricking Arnoldi into running in the orthogonal complement of the dangerous direction.

Figure 7.1 demonstrates this restart strategy in action. As we saw earlier, 25 iterations of shift-and-invert Arnoldi lead to stagnation in 9 of the 10 target eigenpairs. However, we can resolve all 10 target eigenpairs to unit round-off accuracy in 25 iterations if we restart with the Ritz vector corresponding to the dangerous direction after the second iteration. The right Ritz vector is easy to identify: it is most closely aligned with the second iteration y_2. It is worth noting that the Ritz restart strategy seems to be equally successful when eigenpairs are extracted from the Hessenberg matrix H_k instead of $Q_k^* A Q_k$ (see the right panel in Figure 7.1).

8. Multiple dangerous eigenvalues. For simplicity, our analysis has focused on the case where there is just one dangerous eigenvalue. However, other situations may arise more naturally in practice. When eigenvalues are heavily clustered, many dangerous eigenvalues may surround a single pole at various distances. The main message of our results carries over to these cases. To illustrate this, we generate

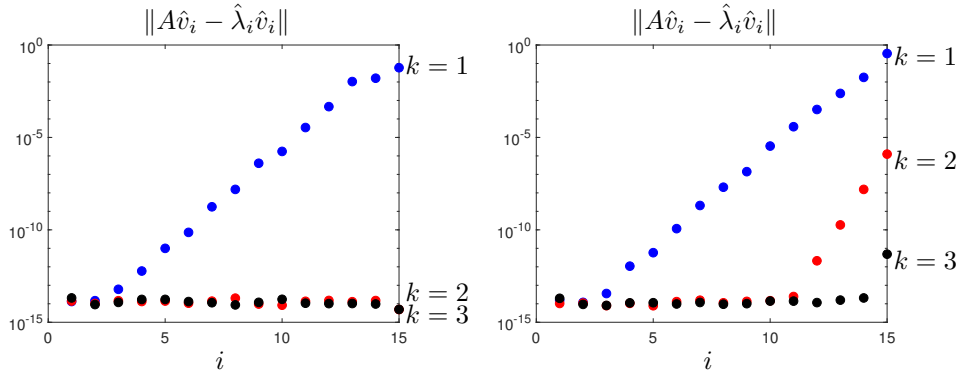

FIG. 8.1. *Convergence with multiple dangerous eigenvalues. On the left, two iterations of rational subspace iteration with a high-quality filter ($\ell = 32$ poles) reduce the residuals of* 15 *target eigenpairs to the order of u, despite exponential clustering of the target eigenvalues at a pole. On the right, three iterations of rational subspace iteration with a medium-quality filter ($\ell = 8$) reduce the residuals of* 15 *target eigenpairs geometrically (see Theorem* 5.4*), with no observable interference from the exponentially clustered eigenvalues.*

a 200×200 symmetric matrix with 15 target eigenvalues in $[10, 15]$, and employ a filter with equally spaced poles on a circular contour centered at 12.5. There are two dangerous eigenvalues at $10+10^{-13}$, and the other 13 target eigenvalues are clustered exponentially at the pole, taking the values $10 + 10^{-i}, i = 0, 1, 2, \ldots, 12$. Thus, there are two dangerous eigenvalues, along with many less harmful but still dangerous eigenvalues. Figure 8.1 shows the results with two rational filters: one excellent and one of medium quality. Just as in sections 3 and 4, we see that twice is enough if the filter quality is high; with a poorer filter, the iterates beyond the second behave as if there was no dangerous eigenvalue (also, see the right panels in Figures 1.1 and 1.2).

Conclusions. Subspace and Arnoldi iterations can be extremely efficient and flexible tools for computing a few target eigenpairs when accelerated with a rational filter, but one must be cautious about eigenvalues near the poles. The damage incurred by such dangerous eigenvalues is confined to the first iteration of subspace iteration: subsequent iterations self-correct, and the eigenpairs are computed to machine precision as orthogonalization effectively deflates the dangerous direction. If the matrix is real-symmetric or, more generally, normal, then the influence of the dangerous eigenvalue is corrected in just two iterations. For matrices whose eigenvectors are not orthogonal (or very close to orthogonal), self-correction occurs geometrically over a series of iterations at a rate of roughly u/d in the best case (it is possible that nonnormal effects cause instability in the worst case). For Arnoldi and similar Krylov schemes, we recommend restarting the iteration with the Ritz approximation to the dangerous eigenvector in order to resolve all target eigenpairs to full precision.

Acknowledgments. We would like to thank Alex Townsend for encouraging us to investigate the stability of contour integral eigensolvers when an eigenvalue is near a quadrature node, as well as for his careful reading of an early draft. We would also like to thank the two anonymous referees for their helpful comments and suggestions which helped us clarify and improve this manuscript.

REFERENCES

[1] A. P. Austin and L. N. Trefethen, *Computing eigenvalues of real symmetric matrices with rational filters in real arithmetic*, SIAM J. Sci. Comput., 37 (2015), pp. A1365–A1387, https://doi.org/10.1137/140984129.

[2] A. Bjorck and G. H. Golub, *Numerical methods for computing angles between linear subspaces*, Math. Comp., 27 (1973), pp. 579–594.

[3] K. R. Davidson and S. J. Szarek, *Local operator theory, random matrices and Banach spaces*, in Handbook of the Geometry of Banach Spaces, Vol. 1, North-Holland, 2001, pp. 317–366.

[4] S. Güttel, E. Polizzi, P. T. P. Tang, and G. Viaud, *Zolotarev quadrature rules and load balancing for the FEAST eigensolver*, SIAM J. Sci. Comput., 37 (2015), pp. A2100–A2122, https://doi.org/10.1137/140980090.

[5] N. J. Higham, *Accuracy and Stability of Numerical Algorithms*, SIAM, 2002, https://doi.org/10.1137/1.9780898718027.

[6] M. H. Holmes, *Introduction to Perturbation Methods*, Texts Appl. Math. 20, Springer, 2013.

[7] A. Horning and A. Townsend, *FEAST for differential eigenvalue problems*, SIAM J. Numer. Anal., 58 (2020), pp. 1239–1262, https://doi.org/10.1137/19M1238708.

[8] J. Kestyn, V. Kalantzis, E. Polizzi, and Y. Saad, *PFEAST: A high performance sparse eigenvalue solver using distributed-memory linear solvers*, in SC'16: Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis, IEEE, 2016, pp. 178–189.

[9] B. N. Parlett, *The Symmetric Eigenvalue Problem*, Classics Appl. Math. 20, SIAM, 1998, https://doi.org/10.1137/1.9781611971163.

[10] G. Peters and J. H. Wilkinson, *Inverse iteration, ill-conditioned equations and Newton's method*, SIAM Rev., 21 (1979), pp. 339–360, https://doi.org/10.1137/1021052.

[11] E. Polizzi, *Density-matrix-based algorithm for solving eigenvalue problems*, Phys. Rev. B, 79 (2009), 115112.

[12] Y. Saad, *Numerical Methods for Large Eigenvalue Problems*, rev. ed., Classics Appl. Math. 66, SIAM, 2011, https://doi.org/10.1137/1.9781611970739.

[13] Y. Saad, *Analysis of subspace iteration for eigenvalue problems with evolving matrices*, SIAM J. Matrix Anal. Appl., 37 (2016), pp. 103–122, https://doi.org/10.1137/141002037.

[14] T. Sakurai and H. Sugiura, *A projection method for generalized eigenvalue problems using numerical integration*, J. Comput. Appl. Math., 159 (2003), pp. 119–128.

[15] G. W. Stewart, *Matrix Algorithms: Volume* II: *Eigensystems*, SIAM, 2001, https://doi.org/10.1137/1.9780898718058.

[16] G. W. Stewart and J.-G. Sun, *Matrix Perturbation Theory*, Elsevier, 1990.

[17] P. T. P. Tang and E. Polizzi, *FEAST as a subspace iteration eigensolver accelerated by approximate spectral projection*, SIAM J. Matrix Anal. Appl., 35 (2014), pp. 354–390, https://doi.org/10.1137/13090866X.

[18] T. Tao, *Nonlinear Dispersive Equations: Local and Global Analysis*, CBMS Reg. Conf. Ser. Math. 106, AMS, 2006.

[19] R. C. Thompson, *The behavior of eigenvalues and singular values under perturbations of restricted rank*, Linear Algebra Appl., 13 (1976), pp. 69–78.

[20] L. N. Trefethen and D. Bau III, *Numerical Linear Algebra*, SIAM, 1997.

[21] L. N. Trefethen and M. Embree, *Spectra and Pseudospectra: The Behavior of Nonnormal Matrices and Operators*, Princeton University Press, 2005.

[22] J. H. Wilkinson, *The Algebraic Eigenvalue Problem*, Numer. Math. Sci. Comput., Oxford University Press, 1965.

[23] P. Zhu and A. V. Knyazev, *Angles between subspaces and their tangents*, J. Numer. Math., 21 (2013), pp. 325–340.